

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 27

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 593 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(i i i) IMMEDIATE SOURCE:

(B) CLONE: Human Fibrinogen A-alpha chain

(i x) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: join(31..84, 1154..1279, 1739..1922, 3055..3200, 3786..5210)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCTAGGAGC	CAOCCCCACC	CYTAOAAAAA	ATO TTT TCC	ATO AGG ATC	OTC TOC	54
			Met Phe Ser Met Arg Ile Val Cys			
			1	5		
CTA GTT CTA	AGT GTO GTO	GGC ACA GCA	TGG GTATGGCCCT	TTTCATTTTT		104
Leu Val Leu	Ser Val Val	Gly Thr Ala	Trp			
10	15					
TCTTCTTGGT	TTCTCTCTGG	TGTTTATTC	ACAAAAGAGCC	TGAGGATCA	AGTCTACCTG	164
CTCTATGTCC	TOACACACTC	TTAGCTTTAT	GACCCCAAGC	CTGGAGGAAA	ATTTCCTGGG	224

66570-334360

-continued

TGGGCTTGAC	ACCTCAAAGAA	TACAGGGTAA	TATGACACCA	AGAGGAAAGAT	CTTAGATGGA	284		
TGAGAGTGT	CAACTACAA	GGAAACTTTA	GCATCTGTCA	TTCAGTCTTA	CCACATTTT	344		
TTTTGTTTT	TTTTAAAAA	GGCAAGAAAT	ATTTGCCATC	CTTGACCTA	TAAAAGCCTT	404		
GTGCATTATA	ATGCTAGTTA	ATGGAATAAA	ACATTTTATG	GTAAAGATTG	TTTTCTTTAG	464		
TTATTAATTT	CTTGCTACTT	GTCCATAATA	AGCAAGAACTT	TTAGTGTTAG	TACAGTTTTT	524		
CTGAAAGGTT	ATTGTTGTGT	TTGTCAAAGAC	AGAAAGAAAAA	GCAAAACGAAT	TATCTTTTGA	584		
AAATATCTTT	CAGTATCAGA	AGAGATTAGT	TAGTAAAGCA	ATACGCTTTT	CCGCAAGTAAT	644		
GGTATTCCTT	TAAATTATGA	ATCCATCTCT	AAAGGTTTACA	TAGAAACTTG	AAAGAGAGAG	704		
GAACATTCA	TTAAGATAGT	CTAGGTTTTT	CTACTGAAAG	AGCAATTACA	GGAGAAAAAG	764		
CTCTACAGTA	GTTCCTCACT	TTCTGTCTGC	AGTCATTAGT	AAAAATGAAA	AGGTAAAAAT	824		
TAACTGATTT	TATAGATTCA	AATAATTTTC	CTTTTAGGAT	GGATTCCTTA	AAACTCCTAA	884		
TATTTATCAA	ATGCTTATTT	AAGTGTACACA	CACAGTTAAG	AAATTTGTAC	ACCTTGTCTC	944		
CTTTAATTC	CATAACAAC	CCATAAAATG	GTCTCTAGGA	TTTCCATTTG	AAAGATAAGAA	1004		
ACCTGAAAG	TGCGAAAGCC	CTGTGTCTGC	TCTCTTAAAT	CTCTGTGAGA	GTGCGATCTC	1064		
TTCTGAGGGA	CTGTGAGGCA	TGCGACTGTC	TCTCTTCTG	GCTAACATTG	CTGTGTCTCT	1124		
CTTTGTGTA	TGTGAATGAA	TCTTTAAAG	ACT GCA GAT	AGT GGT GAA	GGT GAC	1177		
Thr Ala Asp Ser Gly Glu Gly Asp 20 25								
TTT CTA GCT	GAA GGA GGA	GGC GTC	COT GOC	CCA AGG	GTT GTG	GAA AGA	1225	
Phe Leu Ala	Glu Gly Gly Gly	Val Arg Gly	Pro Arg Val	Val Glu Arg				
30 35 40								
CAT CAA TCT	GCC TGC	AAA GAT	TCA GAC	TGG CCC	TTC TGC	TCT GAT	GAA	1273
His Glu Ser	Ala Cys Lys	Asp Ser Asp	Trp Pro Phe	Cys Ser Asp	Glu			
45 50 55								
GAC TGG	GTAAAGAGTC	AGCGGGGGA	GCAAGAGATT	CCITCCCTCT	GATGCTAGAG		1329	
Asp Trp	60							
GGGCTCACAG	GCTGACCTGA	TTGGTCCCA	AACTTTTTT	AAATAGAAAA	TAATTGAATA		1389	
GTACCTTACA	TAGCAAAATA	AGAAAAAGAA	CCTACTCCCA	AGAGCACTGT	TTATTTACCT		1449	
CCCCAACTCT	GGATCATTAG	TGGGTGAACA	GACAGGATT	CAGTTGCATG	CTCAGGCAAA		1509	
ACCAAGCTCC	TGAGTATTGT	GGCTCAATT	TCTGGCACC	TATTTATGCG	TAAAGGACC		1569	
CTCATTCAG	AGTTTCTCTG	CGACCTCTAA	CTAGTCTCT	TACCTACTTT	TAAGCCAACT		1629	
TATCTGGAAG	AGAAAAAGTA	GGAAAGAAATG	GGGCTGCAT	GGAAACATGC	AAAATTATTC		1689	
TGAATCTGAG	AGATAGATCC	TTACTGTAAT	TTTCTCCCTT	CACTTTCAAG	AAC TAC		1744	
Asp Tyr								
AAA TGC	CCT TCT	GOC TGC	AGG ATG	AAA GGG	TTC ATT	GAT GAA	GTC AAT	1792
Lys Cys Pro	Ser Gly Cys	Arg Met Lys	Gly Leu Ile	Asp Glu Val	Asn			
65 70 75								
CAA GAT	TTT ACA	AAC AGA	ATA AAT	AGG CTC	AAA AAT	TCA CTA	TTT GAA	1840
Glu Asp Phe	Thr Asn Arg	Ile Asn Lys	Leu Lys Asn	Ser Leu Phe	Glu			
80 85 90								
TAT CAG	AAO AAC	AAT AAO	GAT TCT	CAT TCG	TTC ACC	ACT AAT	ATA ATG	1888
Tyr Glu Lys	Asn Asn Lys	Ile Asp Ser	His Ser Thr	Thr Thr Asn	Ile Met			
95 100 105 110								
GAA ATT	TTC AGA	GGC GAT	TTT TCC	TCA GGC	AAT A	GTAAATATTA		1932
Glu Ile Leu	Arg Gly Asp	Phe Ser Ser	Ala Asn					
115 120								
CATATTTACT	TCTTTGACTT	TATAACAGAA	ACAACAAAAA	TCCTAAATAA	ATATGATATC		1992	

GATTCGAGGAGG

-continued

CGCTTATATC	TATGACAAAT	TCATCCCAAA	GTACTTAGTG	TAGAAACACA	TACCTTCATA	2052
ATATCCCTGA	AAATTTTAA	AGGAGGCTTT	TGTTTTCTGT	ATTTTTTCAA	AGTAAAAAGAT	2112
GTAACTGAG	ATTGTTTAA	GTACACAAAT	AAGTCAAGAT	TTTGGAATTAA	AACAAAGAAAT	2172
TAAATGTGTT	CTTTTCAACA	GTATATACTG	AAAATAAGAT	GGGTCAAGACT	CTTTGAGTTG	2232
ATATTTTTGT	TTCTGCTTTG	TAAAAGGTGAA	AACTGAGAGG	TCAAAGAACT	TGTTCAAAGGA	2292
CACAGAGCTG	GGAATTC AAC	TCCCAAGACTC	CAGTGAAGCTG	ATTAAGGTAGA	TTTTTAAATTT	2352
TAAAAATATA	GGTCAAAGCTA	CGTCATTCTC	ACAAGTCTACT	CATTAAGGGTT	AGGAAACATT	2412
GCATTCACCTC	TGGGCATGGA	CAAGCAGGTCT	AGGGAGGTCTT	CAGTTTCTCA	AGTTTTGCTT	2472
TGCTTTTTTA	CACCTTCACA	AACACTTGAC	ATTTAAAAATC	AGTGATGCCA	ACACTAGCTG	2532
GCAAAGTGAGT	GATCCTGTTG	ACCCAAAACA	GCTTAGGGAAC	CATTTCAAAT	CTATAGAGTT	2592
AAAAAGAAAA	GCTCATCAGT	AAGAAAAATCC	AATATGTTCA	AGTCCCTTGA	TTAAGGATGT	2652
TATAAAATAA	TTGAAATGCA	ATCAAACCAA	CTATTTTAAAC	TCCAAATTAC	ACCTTTAAAA	2712
TTCCAAAGAA	AGTTCTTCTT	CTATATTTCT	TTGGGATTAC	TAATTOCTAT	TAGGACATCT	2772
TAACTGGCAT	TGATGGAAGG	CTGCAAGGCA	TAAACATTATC	CAAAAAGTCAA	ATGCCCCATA	2832
GGTTTTGAAAC	TCACAGATTA	AACTGTAAAC	AAAAATAAAAT	TAGGCATATT	TACAAAGCTAG	2892
TTTCTTTCTT	TCTTTTTTCT	CTTCTTTTCT	TTCTTTCTTT	CTTCTTTTCT	TTCTTTCTTT	2952
CTTCTTTTCT	TTCTGCTTCC	TTCTTTTCTT	CCTTTCTTTT	TTGCTGGCAA	TTACAGACAA	3012
ATCACTCAAG	AGCTACTTCA	ATAACCATAT	TTTCAATTTT	AG AC CGT	GAT AAT	3065

Ass Arg Asp Asn
125

ACC TAC AAC CGA GTG TCA GAG GAT CTO AGA AOC AGA ATT GAA CTC CTO 3113
Thr Tyr Asn Arg Val Ser Glu Asp Leu Arg Ser Arg Ile Glu Val Leu
130 135 140

AAG CGC AAA GTC ATA GAA AAA GTA CAG CAT ATC CAG CTT CTO CAG AAA 3161
 Lys Arg Lys Val Ile Glu Lys Val Glu His Ile Glu Leu Leu Glu Lys
 145 150 155

AAT GTT AGA GCT CAG TTT GTT GAT ATO AAA CGA CTG GAG GTAACTATGT 3210
Asn Val Arg Ala Glu Leu Val Asp Met Lys Arg Leu Glu
160 165 170

00CT0T00TC	CC0A0T0TCC	TT0TTTTT0A	0TA0A000AA	AA00AA00C0	ATA0TTAT0C	3270
ACT0A0T0TC	TACTATAT0C	AG0A0AAA0T	0TTATATCCA	TCATCTACCT	AAAA0TAG0T	3330
ATTATTTTCC	TCACTCCACA	0TT0AA0AAA	AAAAAATTCA	0A0ATATTAA	0TAAATTTTC	3390
CAAC0TACAT	AGATA0TAA	TCAAA0CAAT	0TTCA0TCCC	T0TCTATTCC	AA0CCATTAC	3450
ATCACCACAC	CTCT0A0CCC	TCA0CCT0AG	TTCACCAA00	ATCATTTAAT	TAGC0TTTCC	3510
TTT0A0A000	AATA0CACCT	TACTCTT0AT	CCATTCT0AG	0CTAA0AT0A	ATTAAACA0C	3570
ATCCATT0CT	TATCCT00CT	A0CCCT0CAA	TACCCAACAT	CTCTTCCACT	0A000T0CTC	3630
0ATA00CA0A	AAACA0A0AA	TATTAA0T00	TAG0TCTCC0	A0TCAAAAAA	AAT0AAACCA	3690
0TTTCCA0AA	00AAAAATTAA	CTACCAG0AA	CTCAATAGAC	0TA0TTTAT0	TATTT0TATC	3750
TACATTTTCT	CTTTATTTTT	CTCCCTCTC	TCTAG 0T0	0AC ATT 0AT	ATT AAG	3803

Val Asp Ile Asp Ile Lys
175

ATC	CGA	TCT	TGT	CGA	GGG	TCA	TGC	AGT	AGG	GCT	TTA	GCT	COT	GAA	GTA	3851
Ile	Arg	Ser	Cys	Arg	Gly	Ser	Cys	Ser	Arg	Ala	Leu	Ala	Arg	Glu	Val	
			180					185					190			

0AT CTG AAO GAC TAT OAA OAT CAG CAG AAO CAA CTT OAA CAG GTC ATT. 3899

-continued

Asp	Leu	Lys	Asp	Tyr	Glu	Asp	Glu	Lys	Glu	Leu	Glu	Val	Ile		
195						200				205					
GCC	AAA	GAC	TTA	CTT	CCC	TCT	AGA	GAT	AGG	CAA	CAC	TTA	CCA	CTG	ATA
Ala	Lys	Asp	Leu	Leu	Pro	Ser	Arg	Asp	Arg	Glu	His	Leu	Pro	Leu	Ile
210						215					220				3947
AAA	ATG	AAA	CCA	GTT	CCA	GAC	TTG	GTT	CCC	GGA	AAT	TTT	AAG	AGC	CAG
Lys	Met	Lys	Pro	Val	Pro	Asp	Leu	Val	Pro	Gly	Asn	Phe	Lys	Ser	Glu
225					230					235					3995
CTT	CAG	AAG	ATA	CCC	CCA	GAG	TGG	AAG	GCA	TTA	ACA	GAC	ATG	CCG	CAG
Leu	Glu	Lys	Val	Pro	Pro	Glu	Trp	Lys	Ala	Leu	Thr	Asp	Met	Pro	Glu
				245					250					255	4043
ATG	AGA	ATG	GAG	TTA	GAG	AGA	CCT	GGT	GGA	AAT	GAG	ATT	ACT	CGA	GGA
Met	Arg	Met	Glu	Leu	Glu	Arg	Pro	Gly	Gly	Asn	Glu	Ile	Thr	Arg	Gly
			260					265					270		4091
GGC	TCC	ACC	TCT	TAT	GGA	ACC	GGA	TCA	GAG	ACG	GAA	AGC	CCC	AGG	AAC
Gly	Ser	Thr	Ser	Tyr	Gly	Thr	Gly	Ser	Glu	Thr	Glu	Ser	Pro	Arg	Asn
		275					280					285			4139
CCT	AGC	AGT	GCT	GGA	AGC	TGG	AAC	TCT	GGG	AGC	TCT	GGA	CCT	GGA	AGT
Pro	Ser	Ser	Ala	Gly	Ser	Trp	Asn	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Ser
	290					295					300				4187
ACT	GGA	AAC	CGA	AAC	CCT	GGG	AGC	TCT	GGG	ACT	GGA	GGG	ACT	GCA	ACC
Thr	Gly	Asn	Arg	Asn	Pro	Gly	Ser	Ser	Gly	Thr	Gly	Gly	Thr	Ala	Thr
305					310					315					4235
TGG	AAA	CCT	GGG	AGC	TCT	GGA	CCT	GGA	AGT	GCT	GGA	AGC	TGG	AAC	TCT
Trp	Lys	Pro	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Ala	Gly	Ser	Trp	Asn	Ser
				325					330					335	4283
GGG	AGC	TCT	GGA	ACT	GGA	AGT	ACT	GGA	AAC	CAA	AAC	CCT	GGG	AGC	CCT
Gly	Ser	Ser	Gly	Thr	Gly	Ser	Thr	Gly	Asn	Glu	Asn	Pro	Gly	Ser	Pro
			340					345					350		4331
AGA	CCT	GGT	AGT	ACC	GGA	ACC	TGG	AAT	CCT	GGC	AGC	TCT	GAA	CAC	GGA
Arg	Pro	Gly	Ser	Thr	Gly	Thr	Trp	Asn	Pro	Gly	Ser	Ser	Glu	Arg	Gly
		355					360					365			4379
AGT	GCT	GGG	CAC	TGG	ACC	TCT	GAG	AGC	TCT	GTA	TCT	GGT	AGT	ACT	GGA
Ser	Ala	Gly	His	Trp	Thr	Ser	Glu	Ser	Ser	Val	Ser	Gly	Ser	Thr	Gly
	370					375					380				4427
CAA	TGG	CAC	TCT	GAA	TCT	GGA	AGT	TTT	AGG	CCA	GAT	AGC	CCA	GGC	TCT
Glu	Trp	His	Ser	Glu	Ser	Gly	Ser	Phe	Arg	Pro	Asp	Ser	Pro	Gly	Ser
385					390					395					4475
GGG	AAC	GGG	AGG	CCT	AAC	AAC	CCA	GAC	TGG	GGC	ACA	TTT	GAA	GAG	GTO
Gly	Asn	Ala	Arg	Pro	Asn	Asn	Pro	Asp	Trp	Gly	Thr	Phe	Glu	Glu	Val
				405					410					415	4523
TCA	GGA	AAT	GTA	AOT	CCA	GGG	ACA	AGG	AGA	GAG	TAC	CAC	ACA	GAA	AAA
Ser	Gly	Asn	Val	Ser	Pro	Gly	Thr	Arg	Arg	Glu	Tyr	His	Thr	Glu	Lys
			420					425					430		4571
CTG	GTC	ACT	TCT	AAA	GGA	GAT	AAA	GAG	CTC	AGG	ACT	GGT	AAA	GAG	AAG
Leu	Val	Thr	Ser	Lys	Gly	Asp	Lys	Glu	Leu	Arg	Thr	Gly	Lys	Glu	Lys
		435					440					445			4619
GTC	ACC	TCT	GGT	AGC	ACA	ACC	ACC	ACG	CGT	CGT	TCA	TGC	TCT	AAA	ACC
Val	Thr	Ser	Gly	Ser	Thr	Thr	Thr	Thr	Arg	Arg	Ser	Cys	Ser	Lys	Thr
		450				455					460				4667
GTT	ACT	AAG	ACT	GTT	ATT	GGT	CCT	GAT	GGT	CAC	AAA	GAA	GTT	ACC	AAA
Val	Thr	Lys	Thr	Val	Ile	Gly	Pro	Asp	Gly	His	Lys	Glu	Val	Thr	Lys
465					470					475					4715
GAA	GTO	GTO	ACC	TCC	GAA	GAT	GGT	TCT	GAC	TGT	CCC	GAG	GCA	ATG	GAT
Glu	Val	Val	Thr		Glu	Asp	Gly	Ser	Asp	Cys	Pro	Glu	Ala	Met	Asp
				485					490					495	4763
TTA	GGC	ACA	TTG	TCT	GGC	ATA	GGT	ACT	CTG	GAT	GGG	TTC	COC	CAT	AAG
Leu	Gly	Thr	Leu	Ser	Gly	Ile	Gly	Thr	Leu	Asp	Gly	Phe	Arg	His	Arg
			500					505					510		4811
CAC	CCT	GAT	GAA	GCT	GCC	TTC	TTC	GAC	ACT	GCC	TCA	ACT	GGA	AAA	ACA
															4855

-continued

His	Pro	Asp	Glu	Ala	Ala	Phe	Phe	Asp	Thr	Ala	Ser	Thr	Gly	Lys	Thr	
		515					520					525				
TTC	CCA	GGT	TTC	TTC	TCA	CCT	ATG	TTA	GGA	GAG	TTT	OTC	AOT	GAG	ACT	4907
Phe	Pro	Gly	Phe	Phe	Ser	Pro	Met	Leu	Gly	Glu	Phe	Val	Ser	Glu	Thr	
	530					535					540					
GAG	TCT	AGG	GGC	TCA	GAA	TCT	GGC	ATC	TTC	ACA	AAT	ACA	AAO	GAA	TCC	4955
Glu	Ser	Arg	Gly	Ser	Glu	Ser	Gly	Ile	Phe	Thr	Asn	Thr	Lys	Glu	Ser	
545					550				555					560		
AGT	TCT	CAT	CAC	CCT	GGG	ATA	GCT	GAA	TTC	CCT	TCC	CGT	GGT	AAA	TCT	5003
Ser	Ser	His	His	Pro	Gly	Ile	Ala	Glu	Phe	Pro	Ser	Arg	Gly	Lys	Ser	
				565				570						575		
TCA	AGT	TAC	AGC	AAA	CAA	TTT	ACT	AGT	AGC	ACG	AGT	TAC	AAC	AGA	OGA	5051
Ser	Ser	Tyr	Ser	Lys	Gln	Phe	Thr	Ser	Ser	Thr	Ser	Tyr	Asn	Arg	Gly	
			580				585						590			
GAC	TCC	ACA	TTT	GAA	AGC	AAO	AGC	TAT	AAA	ATG	GCA	GAT	GAG	GCC	OGA	5099
Asp	Ser	Thr	Phe	Glu	Ser	Lys	Ser	Tyr	Lys	Met	Ala	Asp	Glu	Ala	Gly	
	595					600					605					
AGT	GAA	GCC	GAT	CAT	GAA	OGA	ACA	CAT	AGC	ACC	AAO	AGA	GGC	CAT	GCT	5147
Ser	Glu	Ala	Asp	His	Glu	Gly	Thr	His	Ser	Thr	Lys	Arg	Gly	His	Ala	
	610				615						620					
AAA	TCT	CGC	CCT	OTC	AGA	GGT	ATC	CAC	ACT	TCT	CCT	TTG	GGG	AAO	CCT	5195
Lys	Ser	Arg	Pro	Val	Arg	Gly	Ile	His	Thr	Ser	Pro	Leu	Gly	Lys	Pro	
625					630					635				640		
TCC	CTG	TCC	CCC	TAGACTAAGT	TAAATATTTC	TGCACAGTGT	TCCCATGCCC									5247
Ser	Leu	Ser	Pro		645											
CCTTGCAATTT	CCTTCCTAAC	TCTCTOTTAC	ACOTCATTGA	AACTACACTT	TTTTGOTCTG											5307
TTTTTGOTCT	AGACTGTAAO	TTCTTTGGGG	GCAAGGOCCTT	TGTCTGTCTC	ATCTCTGTAT											5367
TCCCAAAATGC	CTAACAGTAC	AGAGCCATGA	CTCAATAAAT	ACATGTTAAA	TGATGAATG											5427
AATTCCTCTG	AAACTCTATT	TGAGCTTATT	TAOTCAAATT	CTTTCACTAT	TCAAAOTGTG											5487
TGCTATTAGA	ATTGTCACCC	AACTGATTAA	TCACATTTTT	AGTATGTGTC	TCAOTTGACA											5547
TTTAAGTCAO	GCTAAATACA	AGTTGTGTTA	GTATTAAAGTG	AGCTTAAGTA	CCTGTACTGG											5607
TTACTTGCTA	TTAGTTTGTG	CAAGTAAAAAT	TCCAAATACA	TTTGAAGAAA	ATCCCTTTTG											5667
CAATTTGTAG	GTATAAATAA	CCGCTTATTT	GCATAAAGTTC	TATCCCACTG	TAAGTGCATC											5727
CTTTCCTAT	GGAGGGAGAG	AAAGGAGGAA	GAAAGGAGAG	AAAGGAGGAA	AACAATATTT											5787
GCCTTATTTA	ATCTGAGCCG	TGCCTATCTT	TGTAAAGTTA	AATGAGGATA	ACTTCTTCCA											5847
ACCAAGCTTAA	TTTTTTTTTT	AGACTGTGAT	GATGTCCTCC	AAACACATCC	TTCAGGTACC											5907
CAAAATGGCA	TTTTCATAT	CAAGCTATCC	GGATCC													5943

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 644 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Phe Ser Met Arg Ile Val Cys Leu Val Leu Ser Val Val Gly Thr
 1 5 10 15
 Ala Trp Thr Ala Asp Ser Gly Glu Gly Asp Phe Leu Ala Glu Gly Gly
 20 25 30
 Gly Val Arg Gly Pro Arg Val Val Glu Arg His Glu Ser Ala Cys Lys
 35 40 45

644
 544
 444
 344
 244
 144
 44
 4
 1

-continued

Asp	Ser	Asp	Trp	Pro	Phe	Cys	Ser	Asp	Glu	Asp	Trp	Asa	Tyr	Lys	Cys
50						55					60				
Pro	Ser	Gly	Cys	Arg	Met	Lys	Gly	Leu	Ile	Asp	Glu	Val	Asa	Gln	Asp
65					70					75					80
Phe	Thr	Asa	Arg	Ile	Asa	Lys	Leu	Lys	Asa	Ser	Leu	Phe	Glu	Tyr	Gln
			85						90					95	
Lys	Asa	Asa	Lys	Asp	Ser	His	Ser	Leu	Thr	Thr	Asa	Ile	Met	Glu	Ile
			100					105					110		
Leu	Arg	Gly	Asp	Phe	Ser	Ser	Ala	Asa	Asa	Arg	Asp	Asa	Thr	Tyr	Asa
		115					120					125			
Arg	Val	Ser	Glu	Asp	Leu	Arg	Ser	Arg	Ile	Glu	Val	Leu	Lys	Arg	Lys
	130					135					140				
Val	Ile	Glu	Lys	Val	Gln	His	Ile	Gln	Leu	Leu	Gln	Lys	Asa	Val	Arg
145					150					155					160
Ala	Gln	Leu	Val	Asp	Met	Lys	Arg	Leu	Glu	Val	Asp	Ile	Asp	Ile	Lys
			165						170					175	
Ile	Arg	Ser	Cys	Arg	Gly	Ser	Cys	Ser	Arg	Ala	Leu	Ala	Arg	Glu	Val
			180					185					190		
Asp	Leu	Lys	Asp	Tyr	Glu	Asp	Gln	Gln	Lys	Gln	Leu	Gln	Gln	Val	Ile
	195					200						205			
Ala	Lys	Asp	Leu	Leu	Pro	Ser	Arg	Asp	Arg	Gln	His	Leu	Pro	Leu	Ile
	210					215					220				
Lys	Met	Lys	Pro	Val	Pro	Asp	Leu	Val	Pro	Gly	Asa	Phe	Lys	Ser	Gln
225					230					235					240
Leu	Gln	Lys	Val	Pro	Pro	Gln	Trp	Lys	Ala	Leu	Thr	Asp	Met	Pro	Gln
			245						250					255	
Met	Arg	Met	Glu	Leu	Glu	Arg	Pro	Gly	Gly	Asa	Glu	Ile	Thr	Arg	Gly
		260						265					270		
Gly	Ser	Thr	Ser	Tyr	Gly	Thr	Gly	Ser	Glu	Thr	Gln	Ser	Pro	Arg	Asa
		275					280					285			
Pro	Ser	Ser	Ala	Gly	Ser	Trp	Asa	Ser	Gly	Ser	Ser	Gly	Pro	Gly	Ser
	290					295					300				
Thr	Gly	Asa	Arg	Asa	Pro	Gly	Ser	Ser	Gly	Thr	Gly	Gly	Thr	Ala	Thr
305					310					315					320
Trp	Lys	Pro	Gly	Ser	Ser	Gly	Pro	Gly	Ser	Ala	Gly	Ser	Trp	Asa	Ser
			325						330					335	
Gly	Ser	Ser	Gly	Thr	Gly	Ser	Thr	Gly	Asa	Gln	Asa	Pro	Gly	Ser	Pro
			340					345					350		
Arg	Pro	Gly	Ser	Thr	Gly	Thr	Trp	Asa	Pro	Gly	Ser	Ser	Gln	Arg	Gly
		355					360					365			
Ser	Ala	Gly	His	Trp	Thr	Ser	Gln	Ser	Ser	Val	Ser	Gly	Ser	Thr	Gly
	370					375					380				
Gln	Trp	His	Ser	Gln	Ser	Gly	Ser	Phe	Arg	Pro	Asp	Ser	Pro	Gly	Ser
385					390					395					400
Gly	Asa	Ala	Arg	Pro	Asa	Asa	Pro	Asp	Trp	Gly	Thr	Phe	Gln	Gln	Val
			405						410					415	
Ser	Gly	Asa	Val	Ser	Pro	Gly	Thr	Arg	Arg	Gln	Tyr	His	Thr	Gln	Lys
			420					425					430		
Leu	Val	Thr	Ser	Lys	Gly	Asp	Lys	Gln	Leu	Arg	Thr	Gly	Lys	Gln	Lys
		435					440					445			
Val	Thr	Ser	Gly	Ser	Thr	Thr	Thr	Thr	Arg	Arg	Ser	Cys	Ser	Lys	Thr
	450					455					460				
Val	Thr	Lys	Thr	Val	Ile	Gly	Pro	Asp	Gly	His	Lys	Gln	Val	Thr	Lys
465					470					475					480

65470-65476

-continued

Glu Val Val Thr Ser Glu Asp Gly Ser Asp Cys Pro Glu Ala Met Asp
 485 490 495
 Leu Gly Thr Leu Ser Gly Ile Gly Thr Leu Asp Gly Phe Arg His Arg
 500 505 510
 His Pro Asp Glu Ala Ala Phe Phe Asp Thr Ala Ser Thr Gly Lys Thr
 515 520 525
 Phe Pro Gly Phe Phe Ser Pro Met Leu Gly Glu Phe Val Ser Glu Thr
 530 535 540
 Glu Ser Arg Gly Ser Glu Ser Gly Ile Phe Thr Asn Thr Lys Glu Ser
 545 550 555 560
 Ser Ser His His Pro Gly Ile Ala Glu Phe Pro Ser Arg Gly Lys Ser
 565 570 575
 Ser Ser Tyr Ser Lys Glu Phe Thr Ser Ser Thr Ser Tyr Asn Arg Gly
 580 585 590
 Asp Ser Thr Phe Glu Ser Lys Ser Tyr Lys Met Ala Asp Glu Ala Gly
 595 600 605
 Ser Glu Ala Asp His Glu Gly Thr His Ser Thr Lys Arg Gly His Ala
 610 615 620
 Lys Ser Arg Pro Val Arg Gly Ile His Thr Ser Pro Leu Gly Lys Pro
 625 630 635 640
 Ser Leu Ser Pro

(2) INFORMATION FOR SBQ ID NO3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6878 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

- (B) CLONE: human Striatum B-beta chain

(ix) FEATURE:

- (A) NAME/KEY: misc_RNA
- (B) LOCATION: 1469

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 470_583

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 584_5257

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3253_3449

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 3450_3938

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 3939_4122

(ix) FEATURE:

- (A) NAME/KEY: intron
- (B) LOCATION: 4123_5042

(ix) FEATURE:

- (A) NAME/KEY: exon
- (B) LOCATION: 5043_5270

(ix) FEATURE:

TTAATATAAG ATGTAACATA ATCATATTAT GTGCTTATTT TAATGAAATT AOCATTGCTT 683

ATAOTTATGA	AATGGAATTO	TTAACCTCTG	ACTTATTGTA	TTTAAAGAAAT	TTTTCATAGT	743
ATTTCTTATA	TAAAAACAAA	GTAAATTTCTT	GTTTTTCTAGT	TTATCACCTT	TOTTTTCTTA	803
AGATGAGGAT	GGCTTAAGTA	ATGTAAAGATG	TGTTTTTCTC	ACTTGCTATT	CTGAGTACTG	863
TGATTTTCAT	TTACTTTCTAG	CAATACAGGA	TTACAATTAA	GAGGACAAAG	TCTGAAAAATC	923
TCACAAACTA	TAAAAATAATA	AAAGAGCAGA	ATTTTAAAGAT	AAAAOAAACT	GGTGGTAGGT	983
AGATTGTTCT	TTGGTGAAGG	AAAGTAATAT	ATATTGTTAC	TGAGATTACT	ATTTATAAAA	1043
ATTATAACTA	AOCCTAAAAG	CAAAATACAT	CAAOTOTAAT	GATAAGAAAAT	GAAATATTGC	1103
TTTTTTTCA	TGAAAAATTG	AAATTAGAGT	TAGTGTGTAT	TGTTATTATT	AATAOTTATG	1163
AAACACGGTT	CAOTCTAATT	TATTTATTTG	TAGAACAOTT	TGTCTCAAC	TATTATTTTT	1223
GCTGACTTAT	TGCTOTTAAT	TTGAGTTAC	TAAAAATACA	GAAATGCATT	TAGGACAAATG	1283
GATATTTAAG	AAATTTAAAT	TTTATCATCA	AACGTATCAT	GGCCAAATTT	CTTACATATA	1343
GCATAGTATC	ATTAAACTAG	AAATAAGAAAT	ACACAATAAT	ATTTAAATGA	AGTGATTTCAT	1403
TTGGGATCAT	TATTGAGTTT	CAAGGGAAC	TGAGTGTGTT	ACTTATCAAG	CTCTACATGT	1463
AAAGACATAT	AGTTAATCTG	GTGTGTGTGT	TAAAAACATA	TGTTAATCT	GGTTAAOTCT	1523
GGTTAATCAT	ATTAGGTAAAG	AAAAATGTAA	AGAAATGTGT	AGACGAAATTT	TTTGTAAAOT	1583
ACTCTGCAAA	GCACTTTCAC	ATTTCTGCTT	ATCAACTAAA	CCTCACAGAG	ATAOTTAAAT	1643
AGTTTAAGCT	TTAAAAATGA	TTTTGATTAT	TCAACAAAGT	GCCTTCATAA	TTTCTTTAAG	1703
TGTTTTTCTT	TAAGTATATA	CTTTCTTTAA	ATATTTTTTA	AAATTTCTTT	TTCTCTAGTA	1763
AAGCCAGACC	ATCCATGCTA	CCTCTCTAGT	GGCAGCTCTG	AAATAAAAAAG	AAATAOTTTT	1823
CTCTGTTATA	ATTGTATTTG	TAATAAGCAG	ATGAATCACA	TTTCTTAAAA	TTTGTTTTTAG	1883
AGAGGGTAAAG	CTCTGACTAG	GACCATGACT	TCAATGTGAA	ATATGTATAT	ATCCTCCGAA	1943
TCTTTACATA	TTAAGAAATGT	ATATAGTCAA	CTGGTTAAAC	AGGAAAAATCT	GGAACAGCCT	2003
GGCTGGGTTT	TAATCTTAAG	ACCATCTTAC	TAAATGTTAA	ATAATATTAT	AATCTAATGA	2063
ATAAATGACA	ATGCAATTCC	AAATAAGAGTT	CATCTGATGA	CTTCTAGACT	CACAAAAATTO	2123
CAAGAGAGCT	CAOTTGTTGC	TCAOTTGTTT	CAAAATCATGT	CGTTTGTTAA	TTTGTAAATTA	2183
AGCTCCAAAG	GATGTATAGC	TACTGACAAA	AAAAAAAATG	AGAAATGTAGT	TAATCCAAAT	2243
CAAAACITTC	CTATTGCAAT	GCATATTTTC	TGCTTCATTA	TCCTTTAATA	TAATATTTTA	2303
AGTTAGCAAG	TAATTTTAAAT	TACAATOCAC	AAGCCTTGAG	AATTATTTTA	AATATAAGAA	2363
AATCATAATG	TTTGATAAAG	AAATCATGTA	AGAAATTTCA	AGATAATGTT	TTAACAAATA	2423
ATTTTGTGTA	TAGAAAGATA	GACTAAAAAT	GAAATTCGAA	GTGGAGAGGA	CAGTTAAACT	2483
GTAGTACTTG	TTATGTGTGA	TTCCAATAAA	AATAGTAAATG	AGCAGTTATT	ATTGCCAAGT	2543
ACTGTTCTGA	GGGTACCAT	TGCAATAAAT	TATTTAATCC	TTACAATAAT	CTTGTAAAGG	2603
AGATTCAAAAC	TATCATTACA	CTTATTTTAC	AGATGAGAAA	ACTGGGGCAC	AGATAAAGCA	2663
ACTTGCCCAA	GGTCTCATAG	CTGTAAATCA	ACCCTACGTT	CAAGACCTAC	AAGTAGCCGA	2723
GCTCCAGAGT	ACATTATGAG	GGTCAAAAGAT	TGTCTTATTA	CAAAATAAAT	CCAAATAGAA	2783
TCAACCTTTA	ATAAGTCTTT	AATGTCTCTT	AAATATGTTT	ATATAAGAGT	CTAATCAGCA	2843
ATTCACAAAA	ATGAAAAGTAG	GGAAATGATT	AACAATAATC	ATAGGAAATCT	AACAATCCAA	2903
GTGGCTTGAAG	AATATTTCATT	CTTCTTGAAC	GTATAAGATT	TTTACAATTT	CGTAAATTTCC	2963
AATGTATGTT	TTAGGAATAT	GAGGTCTATTA	CTATTCTATA	TCTGATACAG	CTTTATCTTA	3023
AGGCCTCTCT	TTAAAAACTA	CAGTGCATCA	TAGCTTTTTT	GTGCAAGTTGG	TCITTTCTACT	3083

-continued

OTTACTGAAC	AOTAAOCAAC	CTACAGATT	ACTATCACCA	ACCAOCCAOT	TOATGATCT	3143
TAAOCAAAAT	ATCAAOCCTT	TOATAACCTA	AATTATAAAA	TGAOOGTOTT	OGAATAOTTA	3203
CATTCCAAAT	CTTCTATAAC	ACTCTGTATT	ATATTTCTOC	CTCATTCCCT	GTAG GGT	3260
					Gly	
TTC TTC AGT OCC COT GGT CAT COA CCC CTT GAC AAG AAG AGA GAA GAG	3308					
Phe Phe Ser Ala Arg Gly His Arg Pro Leu Asp Lys Lys Arg Glu Glu						
40 45 50 55						
GCT CCC AOC CTO AOO CCT OCC CCA CCG CCC ATC AGT GGA GGT GGC TAT	3356					
Ala Pro Ser Leu Arg Pro Ala Pro Pro Pro Ile Ser Gly Gly Gly Tyr						
60 65 70						
COG GCT COT CCA OCC AAA OCA OCT OCC ACT CAA AAG AAA GTA GAA AGA	3404					
Arg Ala Arg Pro Ala Lys Ala Ala Ala Thr Glu Lys Lys Val Glu Arg						
75 80 85						
AAA GCC CCT GAT OCT GGA GGC TOT CTT CAC GCT GAC CCA GAC CTO	3449					
Lys Ala Pro Asp Ala Gly Gly Cys Leu His Ala Asp Pro Asp Leu						
90 95 100						
GTGGGTGCAC TOATGTTTCT TCGAGTGGTG GCTCTCTCAT GCAGAGAAAAG CCTGTAGTCA	3509					
TGGCAGTCTG CTAATGTTTC ACTGACCCAC ATTACCATCA CTGTTATTTT GTTTGTTTAT	3569					
TTTGGAAATA AAATTCAAAA CATAAACATA TTGGGCCTTT GGTTTAAGCT TTCTTTCTTG	3629					
TTTTCTTTGG TCTGGGCCCA AAATTTCAAA TTAAGATATG TGGGTGCCAC CTTTCCATTT	3689					
GTATTTTCCC ACTGCCTTTG TTTAGTTGGT AAAATTTTCA TAGCCCAATT ATATTTTTC	3749					
TGGGTAAAGT AATATTTTAA ATCTCTATGA GAGTATGATG ATGACTTTTC AATTTCTGOT	3809					
CTTACAGAAA ACCAAATAAT AAATTTTAT GTTGGCTAAT GGTATCGCTG AATTTTCCTA	3869					
TGTGCTATTT TAACAAATGT CCATGACCCA AATCCTTCAT CTAATGCTG CTATTTTCTT	3929					
TOTTTTTAG GGG GTO TTO TOT CCT ACA GGA TOT CAG TTO CAA GAG OCT	3977					
Gly Val Leu Cys Pro Thr Gly Cys Glu Leu Glu Glu Ala						
105 110 115						
TTO CTA CAA CAG GAA AOO CCA ATC AGA AAT AGT GTT GAT GAG TTA AAT	4025					
Leu Leu Glu Glu Glu Arg Pro Ile Arg Asn Ser Val Asp Glu Leu Asn						
120 125 130						
AAC AAT GTO GAA GCT GTT TCC CAG ACC TCC TCT TCT TCC TTT CAG TAC	4073					
Asn Asn Val Glu Ala Val Ser Glu Thr Ser Ser Ser Ser Ser Phe Glu Tyr						
135 140 145						
ATG TAT TTO CTO AAA GAC CTO TGG CAA AAG AGO CAG AAG CAA GTA AAA	4122					
Met Tyr Leu Leu Lys Asp Leu Trp Glu Lys Arg Glu Lys Glu Val Lys						
150 155 160						
GTAGATATCC TTGTGCTTTC CATTGATTT TCACTATATA AATTGGAACC GTTAGACTOC	4182					
CACGAGAAAT CATGGTTGTG AGAAGATTAA CATTCTGTGG TTAOTGAATA GCATTCATAC	4242					
GCTTTTGGGC ACCTTCCCTT GCAACTTGGC AGATAAGCAC TATTCACTC TTATTCCCAO	4302					
TCTGACATCA GCAAGTGTGA TTTTCTATGA AAAATTCTAC TATGACTCCT TATTTTAAAT	4362					
ATACAAAGAAA CTTGTGACTC AGAAGATAAT ATTTACAGAG TGGAAAAAAA CCCCTAGCAT	4422					
TTATAOTTTT AACATTTOAG GTTTTGAATG AGAGAGTTAT CCATAATATA TTCAATTOTO	4482					
TTOTGGATAA TGACACCTAA CCTGTGAATC TTGAAGTCAO AATGTTGAGT GCTGTTGACT	4542					
TGGTGGTCAO GAAACAOGTA GTGCGTGAOC CTGGCACAAG CATCTCAOTG AGTAOCATAC	4602					
CCACAOTTTG AAATTTTTC AAGAAATCAA AGGAATCATO ACATCTTATA AATTTCAAAG	4662					
TTCTGCTATA CTTATGTGAA ATGGATAAAT AAATCAAACA TATCCACTCT GTAAAGATTGA	4722					
ACTTCTCAGA TGGAAAGACCC CAATACTGCT TTCTCCTCTT TTCCCTCACC AAAGAAATAA	4782					
ACAACCTATT TCATTTATTA CTGGAACACAA TCTTTAGCGT ATACCTATGG TAAATTACTA	4842					

GCTTQ = 5072260

-continued

GTATGTTGTT TAGGATTTAT GTTAATTTGT ATATGTCATG CCCCCAATCA TTTCCACTAA	4902
ATATGACTAT ATATCATAAC TGCTTGGTGA TAGCTCAATG TTTAATAATT TATTCTCAGA	4962
AAATCAAAAT TOTATAOTTA AATACATTAG TTTTATGAGG CAAAAATGCT AACTATTTCT	5022
ACATAATTTT ATTTTTCAG AT AAT GAA AAT GTA GTC AAT GAG TAC TCC	5071
Asp Asn Glu Asn Val Val Asn Glu Tyr Ser 165 170	
TCA GAA CTG GAA AAG CAC CAA TTA TAT ATA GAT GAG ACT GTG AAT AGC	5119
Ser Glu Leu Glu Lys His Glu Leu Tyr Ile Asp Glu Thr Val Asn Ser	175 180 185
AAT ATC CCA ACT AAC CTT COT GTG CTT COT TCA ATC CTG GAA AAC CTG	5167
Asn Ile Pro Thr Asn Leu Arg Val Leu Arg Ser Ile Leu Glu Asn Leu	190 195 200 205
AGA AGC AAA ATA CAA AAG TTA GAA TCT GAT GTC TCA GCT CAA ATG GAA	5215
Arg Ser Lys Ile Glu Lys Leu Glu Ser Asp Val Ser Ala Glu Met Glu	210 215 220
TAT TGT CCG ACC CCA TCG ACT GTC AGT TCG AAT ATT CCT GTG GTG TCT	5263
Tyr Cys Arg Thr Pro Cys Thr Val Ser Cys Asn Ile Pro Val Val Ser	225 230 235
GGC AAA G GTAAGTATT CATAAACATA TTTTAAAGAA GTTCCAGAAAG AACTCACACA	5320
Gly Lys CCAAAAATAA GAGAACAAACA ACAACAACAA AAATGCTAAG TGGATTTTCC CAACAGATCA	5380
TAATGACATT ACAATACATC ATAAAAATAT CCTTAAGCAAG TTGTGTTTTG GACTGGCCTG	5440
GTGCATTTGC TGGTTTTGAT GAGCAAGGATG GGGCACAGGT AGTCCCAAGGG GTGGCTGATG	5500
TGTGCATCTG COTACTGGCT TGAACAGATG GCAAGAACAC AGATAGATGT AGAAGTTTCT	5560
CCATTTTGTG TTTTCTGGGA GCTCATGGAT ATTCCAAGGAC ACAAAGGGTG GAGAAAGAGCT	5620
TTGTTTATCC TCTTAAGCAAG TAAACGTCCT CAAAAGTGGG TTGGACTTAC TAAAGTAAAA	5680
TGAAAAATCTA ATATTTGTTA TATTATTTTC AAAGGTCTAT AATAACACAC TCCTTAATAA	5740
CTATGTAAT GTTATTTTAA AGAATTTGGT ACTAAATACA AAGTAATTAT GTCATAAACC	5800
CCTGAACATA ATGTTGTCTT ACATTTGCAAG AA TOT GAG GAA ATT ATC AGG AAA	5853
Glu Cys Glu Glu Ile Ile Arg Lys 240 245	
GGA GGT GAA ACA TCT GAA ATG TAT CTC ATT CAA CCT GAC AGT TCT GTC	5901
Gly Gly Glu Thr Ser Glu Met Tyr Leu Ile Glu Pro Asp Ser Ser Val	250 255 260
AAA CCG TAT AGA GTA TAC TOT GAC ATG AAT ACA GAA AAT GGA G	5944
Lys Pro Tyr Arg Val Tyr Cys Asp Met Asn Thr Glu Asn Gly	265 270 275
GTAAAGCTTC GACAGTTGTT GACCTGTTGA TCTGTAAATTA TTTGGATACC GTAAAAATGCC	6004
AGGAAACAAAG GCCAGGTGTG GTGGCTCATA CCTGTAAATC CAGCACCTTG GAGGGCCAAA	6064
GTGGGCTGAT AGCTTGAAGC TAGGAGTTTG AAAGTAAAGC GTGGCAACATA ATGAGACCTT	6124
AACTCTACAA AAAAAAAAAA AATACCAAAA AAAAAAAAAA AATCAAGCTGT GTTGGTAAAT	6184
TGTGGCTGTA GTCCCAAGTA TCCAAGGAGG TAGGATGGGA GATCACCTGA GGGCACAAAC	6244
TGGAGTCTTG ATCATGCTAC TGAAGTGTAG CCTGGGCAAC AGAGGATAGT GAGATCCTGT	6304
CTCAAAAAAAAA AAAAAATTAAT AAAAAAGCAAG GAAACAAAGAC TTAAGCTCTAA CATCTAACAT	6364
AGCTGACAAA GGAAGTAATTT GATGTGGAAAT TCAACCTGAT ATTTAAAAAT TATAAAATAT	6424
CTATAATTCA CAATTTGGGG TAAAGATAAAG CACTTGCAGT TTCCAAAGAT TTTACAAATT	6484
TACCTCTCAT ATTTATTTCC TTATTGTGTC TATTTTAGAG CACCAAAATAT ATACTAAATG	6544
GAATGGACAG GGGATTCAAG TATTATTTTC AAAGTGACAT TATTTGCTGT TGGTTAATAT	6604

665770-6372260

-continued

ATGCTCTTTT	TGTTTCTGTC	AACCAAAAG	GA	TGG	ACA	GTG	ATT	CAO	AAC	COT	6655
Gly Trp Thr Val Ile Glu Asn Arg 280 285											
CAA	GAC	GGT	AGT	GTT	GAC	TTT	GGC	AAG	AAA	TGG	6703
Glu	Asp	Gly	Ser	Val	Asp	Phe	Gly	Arg	Lys	Trp	
290 295 300											
GGA	TTT	GGA	AAT	GTT	GCA	ACC	AAC	ACA	GAT	GGG	6751
Gly	Phe	Gly	Asn	Val	Ala	Thr	Asn	Thr	Asp	Gly	
305 310 315											
CTA	CCA	G	GTAACGAACA	GGCATGCAAA	ATAAAAATCAT	TCTATTTTGA	ATGGGATTTT	6808			
Leu	Pro										
TTTTAATTAA	AAAACATTCA	TTGTTTGAAG	CCTGTTTTAG	GCAGTTAAGA	GGAGTTTCTT	6868					
GACAAAAAATG	TGGAAAGCTAA	AGATAAGGGA	AGAAAAGGCAO	TTTTTAATTT	CCCAAAATTT	6928					
TATTTTTGOT	GAGAGATTTT	ATTTTGTITT	TCTTTTAG	GT	GAA	TAT	TGG	CTT	6980		
Gly Glu Tyr Trp Leu 320											
GGA	AAT	GAT	AAA	ATT	AGC	CAO	CTT	ACC	AAG	ATG	7028
Gly	Asn	Asp	Lys	Ile	Ser	Glu	Leu	Thr	Arg	Met	
325 330 335 340											
TTG	ATA	GAA	ATG	GAG	GAC	TGG	AAA	GGA	GAC	AAA	7076
Leu	Ile	Glu	Met	Glu	Asp	Trp	Lys	Gly	Asp	Lys	
345 350 355											
GGA	GGA	TTC	ACT	GTA	CAO	AAT	GAA	GCC	AAC	AAA	7124
Gly	Gly	Phe	Thr	Val	Glu	Asn	Glu	Ala	Asn	Lys	
360 365 370											
AAC	AAA	TAC	AGA	GGA	ACA	GCC	GGT	AAT	GCC	CTC	7172
Asn	Lys	Tyr	Arg	Gly	Thr	Ala	Gly	Asn	Ala	Leu	
375 380 385											
CAO	CTG	ATG	GGA	GAA	AAC	AAG	ACC	ATG	ACC	ATT	7220
Glu	Leu	Met	Gly	Glu	Asn	Arg	Thr	Met	Thr	Ile	
390 395 400											
TTC	AGC	ACG	TAT	GAC	AGA	GAC	AAT	GAC	GGC	TG	7262
Phe	Ser	Thr	Tyr	Asp	Arg	Asp	Asn	Asp	Gly	Trp	
405 410 415											
CACCTTTTGC	TCCTGCTTTA	AAAAATCACAC	TAATATCATT	ACTCAGAAATC	ATTAACAATA	7322					
TTTTTAATAG	CTACCACTTC	CTGGGCACTT	ACTGTCAGCC	ACTGTCTTAA	GCTCTTTATG	7382					
CATCACTCGA	AAGCATTTC	ACTATAAGGT	AGACATTCTT	ATTCTCATTT	TACAGATGAG	7442					
ATTTAAGAGAG	ATTACGTGAT	TTGTCCAATG	TCACACAACCT	ACCCAGAGAT	AAAACCTAGAA	7502					
TTTGAAGCACA	GTTACTTTCT	GAATAATGAG	CATTTAAGATA	AATACCTATA	TCTCTATATT	7562					
CTAAAGTGTG	TGTGAAAACT	TTCATTTTCA	TTTCCAAGGT	TCTCTGATAC	TAAAGGTTTGT	7622					
AAAAAGCTATT	ATTCCAATAT	AAAATAACAA	ACACAATCCC	TAGATGGATT	GCCACAAAGG	7682					
CCCAATTATC	TCTCTTTCTT	GCTATAAGGC	ACAAGGAGGTC	TTTGGTGTAT	TAGTGTGACT	7742					
CTATGTATAG	CACCCAAAAG	AAAAGACTACT	GTGCACACGA	GTGTAAGCAAT	CTTTTATGGG	7802					
TAATCTGCAA	AACGTAACTT	GACCACCOTA	GTTCTGTTTC	TAATAACGCC	AAACACATTT	7862					
TCTTTCAO	G	TTA	ACA	TCA	GAT	CCC	AGA	AAA	CAO	TOT	7910
Leu Thr Ser Asp Pro Arg Lys Glu Cys Ser Lys Glu Asp 420 425											
GGT	GGT	GGA	TGG	TGG	TAT	AAT	AGA	TGT	CAT	GCA	7958
Gly	Gly	Gly	Trp	Trp	Tyr	Asn	Arg	Cys	His	Ala	
430 435 440											
AGA	TAC	TAC	TGG	GGT	GGA	CAO	TAC	ACC	TGG	GAC	8006
Arg	Tyr	Tyr	Trp	Gly	Gly	Glu	Tyr	Thr	Trp	Asp	
Met Ala Lys His Gly											

GCTCTTTT

-continued

445				450				455				460				
ACA	GAT	GAT	GOT	GTA	GTA	TGG	ATG	AAT	TGG	AAO	GGG	TCA	TGG	TAC	TCA	8054
Thr	Asp	Asp	Gly	Val	Val	Trp	Met	Asn	Trp	Lys	Gly	Ser	Trp	Tyr	Ser	
				465					470					475		
ATG	AOO	AAO	ATO	AOT	ATO	AAO	ATC	AOO	CCC	TTC	TTC	CCA	CAO	CAA	TAOTCCCC	8109
Met	Arg	Lys	Met	Ser	Met	Lys	Ile	Arg	Pro	Phe	Phe	Pro	Gln	Gln		
			480					485					490			
TAC	OTA	GA	TTT	OTCTCTC	TG	TATGTG	AC	AAC	ATTTTT	TG	TAC	ATTATGT	TATT	OGAA	TT	8169
TTCT	TT	CATA	CATT	TATATTC	CTCT	AAAA	ACT	CTCA	AGCAGA	CG	TGAG	TGTGT	ACT	TTTT	TGAA	8229
AAAA	GT	TATA	GATA	AAATTAC	AT	AAAA	TAG	CAC	ATGATT	T	CTTT	TGTTT	TC	TT	CATTTC	8289
TCTT	GT	CTCAC	CCA	AGAA	GT	AC	AAAA	GTAT	AGTTTT	GTACA	GA	TTTG	TGT	TC	ATAATTC	8349
AGTT	CT	AGTT	GATT	OCGAGA	AT	TTTCAA	AT	AA	OGAA	GA	AG	GTCT	TTTT	AT	CCTT	8409
OG	AAAA	ACCAT	GAC	OGAAAA	AG	AAAA	ACTG	AT	GT	TA	AAAA	AGT	CC	ACT	TTTTAA	8469
ATTT	AT	GTAG	GAT	CTGTCAA	AG	AAAA	ACTTC	CA	AAAA	AGATT	TAT	TAAT	TA	AA	ACCA	8529
GT	TC	CAATAA	GT	TAAT	GT	TT	TTTTTT	TG	TA	ATCC	CAC	AC	AT	TC	AATG	8589
CAC	TT	GTAG	GA	OGAGAA	AG	C	TTTCA	CAAC	CT	CAAA	TAGC	TA	AT	AA	ACCG	8649
TTT	GA	AGATT	T	AAAA	CT	GA	CT	TAGG	AC	GG	CAC	GG	TT	GT	TA	8709
ACT	TT	GGGAG	G	CTG	AG	GG	CG	GT	CAC	AA	GG	TC	AG	AC	CA	8769
AT	GT	TA	AA	AC	CC	AT	CT	CTA	AAAA	TAT	AGC	CA	GG	CG	TGGT	8829
T	TA	AG	TT	CCC	TA	GG	CT	CTO	AA	GG	CT	CTO	AA	GG	CT	8889

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 491 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: protein

(x i) SEQUENCE DESCRIPTION: SBQ ID NO:4:

Met 1	Lys	Arg	Met	Val 5	Ser	Trp	Ser	Phe	His 10	Lys	Leu	Lys	Thr	Met 15	Lys
His	Leu	Leu	Leu 20	Leu	Leu	Leu	Cys	Val 25	Phe	Leu	Val	Lys	Ser 30	Gln	Gly
Val	Asn	Asp 35	Asn	Glu	Glu	Gly	Phe 40	Phe	Ser	Ala	Arg	Gly 45	His	Arg	Pro
Leu	Asp 50	Lys	Lys	Arg	Glu	Glu 55	Ala	Pro	Ser	Leu	Arg 60	Pro	Ala	Pro	Pro
Pro 65	Ile	Ser	Gly	Gly	Gly 70	Tyr	Arg	Ala	Arg	Pro 75	Ala	Lys	Ala	Ala	Ala 80
Thr	Gln	Lys	Lys	Val 85	Glu	Arg	Lys	Ala	Pro 90	Asp	Ala	Gly	Gly	Cys 95	Leu
His	Ala	Asp	Pro 100	Asp	Leu	Gly	Val	Leu 105	Cys	Pro	Thr	Gly	Cys 110	Gln	Leu
Gln	Glu	Ala 115	Leu	Leu	Gln	Gln	Glu 120	Arg	Pro	Ile	Arg	Asn 125	Ser	Val	Asp
Gln	Leu 130	Asn	Asn	Asn	Val	Glu 135	Ala	Val	Ser	Gln	Thr 140	Ser	Ser	Ser	Ser
Phe 145	Gln	Tyr	Met	Tyr	Leu 150	Leu	Lys	Asp	Leu	Trp 155	Gln	Lys	Arg	Gln	Lys 160
Gln	Val	Lys	Asp	Asn 165	Glu	Asn	Val	Val	Asn 170	Glu	Tyr	Ser	Ser	Glu 175	Leu

-continued

Glu	Lys	His	Gln	Leu	Tyr	Ile	Asp	Glu	Thr	Val	Asn	Ser	Asn	Ile	Pro
			180					185					190		
Thr	Asn	Leu	Arg	Val	Leu	Arg	Ser	Ile	Leu	Glu	Asn	Leu	Arg	Ser	Lys
		195					200					205			
Ile	Gln	Lys	Leu	Glu	Ser	Asp	Val	Ser	Ala	Gln	Met	Glu	Tyr	Cys	Arg
	210					215					220				
Thr	Pro	Cys	Thr	Val	Ser	Cys	Asn	Ile	Pro	Val	Val	Ser	Gly	Lys	Glu
225					230					235					240
Cys	Glu	Glu	Ile	Ile	Arg	Lys	Gly	Gly	Glu	Thr	Ser	Glu	Met	Tyr	Leu
				245					250					255	
Ile	Gln	Pro	Asp	Ser	Ser	Val	Lys	Pro	Tyr	Arg	Val	Tyr	Cys	Asp	Met
			260					265					270		
Asn	Thr	Glu	Asn	Gly	Gly	Trp	Thr	Val	Ile	Gln	Asn	Arg	Gln	Asp	Gly
		275					280					285			
Ser	Val	Asp	Phe	Gly	Arg	Lys	Trp	Asp	Pro	Tyr	Lys	Gln	Gly	Phe	Gly
	290					295					300				
Asn	Val	Ala	Thr	Asn	Thr	Asp	Gly	Lys	Asn	Tyr	Cys	Gly	Leu	Pro	Gly
305					310					315					320
Glu	Tyr	Trp	Leu	Gly	Asn	Asp	Lys	Ile	Ser	Gln	Leu	Thr	Arg	Met	Gly
				325					330					335	
Pro	Thr	Glu	Leu	Leu	Ile	Glu	Met	Glu	Asp	Trp	Lys	Gly	Asp	Lys	Val
			340					345					350		
Lys	Ala	His	Tyr	Gly	Gly	Phe	Thr	Val	Gln	Asn	Glu	Ala	Asn	Lys	Tyr
		355					360					365			
Gln	Ile	Ser	Val	Asn	Lys	Tyr	Arg	Gly	Thr	Ala	Gly	Asn	Ala	Leu	Met
	370					375					380				
Asp	Gly	Ala	Ser	Gln	Leu	Met	Gly	Glu	Asn	Arg	Thr	Met	Thr	Ile	His
385					390					395					400
Asn	Gly	Met	Phe	Phe	Ser	Thr	Tyr	Asp	Arg	Asp	Asn	Asp	Gly	Trp	Leu
				405					410					415	
Thr	Ser	Asp	Pro	Arg	Lys	Gln	Cys	Ser	Lys	Glu	Asp	Gly	Gly	Gly	Trp
			420					425					430		
Trp	Tyr	Asn	Arg	Cys	His	Ala	Ala	Asn	Pro	Asn	Gly	Arg	Tyr	Tyr	Trp
		435					440					445			
Gly	Gly	Gln	Tyr	Thr	Trp	Asp	Met	Ala	Lys	His	Gly	Thr	Asp	Asp	Gly
	450					455					460				
Val	Val	Trp	Met	Asn	Trp	Lys	Gly	Ser	Trp	Tyr	Ser	Met	Arg	Lys	Met
465					470					475					480
Ser	Met	Lys	Ile	Arg	Pro	Phe	Phe	Pro	Gln	Gln					
				485					490						

(2) INFORMATION FOR SEQ ID NO-5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10364 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: DNA (genomic)

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: human fibrinogen gamma chain

(i x) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
.. 2603, 4211..4341, 4643..4778, 5758..5942, 7426
.. 7703, 9342..9571)

(x i) SEQUENCE DESCRIPTION: SEQ ID NO-3:

CTACACACTT	CTTGAAGGCA	AAGGCAATOC	TGAAGTCACC	TTTCATGTTT	AAATCATATT	60
AAAAAGTTAG	CAAGATGTAA	TTATCAOTOT	ACTATGTAAA	TCTTTGTGAA	TGATCAATAA	120
TTACATATTT	TCATTATATA	TATTTTAGTA	GATAATATTT	ATATACATT	AACATTCTAA	180
ATATAGAAA	TTTACAGAGA	AAAATAAAGC	CTTTTTTTCC	AATCCTGTCC	TCCACCTCTG	240
CATCCCATTC	TTCTTCACAG	AAGCAACTGA	TTCAAGTCAT	TACATAGTTA	TTGAGTGTTA	300
ACTACAACCTA	TGTTAAOTAC	AGCTATATAT	GTTAGATGCC	GTAOCCACA	AAATCAOTTT	360
ACAATCTAAT	GCAOTGGATA	CAOCATGTAT	ACATATAATA	TAAAGTTGCT	ACAAATGCTA	420
TCTGAGGTAG	AGCTGTTTTG	AAOAACTACT	ATACTTAAAT	OTTTAATTCA	ACTOACTTGA	480
TTGACAACCTG	ATTAOCTGAG	TGAAAAAGAT	GGATGAGAAA	GATTGTGAGA	CTTAATTGOC	540
TGTTGTTATG	GTGATATGAT	TGACAATAAC	TGCTAAGTCA	GAGAGGGATA	TATTAAGGAG	600
GAGAAAGAAA	GCAACAAATC	TGTTTTTGAT	GTGTTCACTT	TGTTATAATT	ATTGATTATT	660
TACTGAATAT	GAATATTTAT	CTTTGTTTTT	GAOTCAATAA	ATATACCTTT	GTAAAGACAG	720
AAATTAAGTA	TTAGTATTTT	TTTCAAACTG	GAGGCATTTT	TCCCACTAAC	ATATTTTCATC	780
AAAACTTATA	ATAAGCTTGG	TTCCAAGAGG	AGAAATGAGG	GATAACCAAA	AATAAGAGACA	840
TTAATAATAG	TGTAACGCCC	AGTGATAAAT	CTCAATAGGC	AGTGATGACA	GACATGTTTT	900
CCCAACACACA	AGGATGCTGT	AAAGGGCCAA	CAGAAATGAT	GGCCCTCTCC	CAGCACCTCA	960
TTTTGCCCCCT	TCCTTCAOCT	ATGCTCTTAC	TCTCCTTTAG	ATACAAAGGA	GGTGGAATTT	1020
TCTCTTCTCT	GAGATAGCTT	GATGGAACCA	CAGGAACAA	GAAOTGGGCT	CCTGGCTCTT	1080
TTCTCTGTGG	CAGATGGGGT	GCCATGCCCC	CCTTCAGACA	AAAGGAAAGAT	TGAGCTCAAA	1140
AGCTCCCTGA	GAAOTGAGAG	CCTATGAACA	TGTTTGACAC	AGAGGGAGAG	GAATGTATTT	1200
CCAGGGTCAT	TCATTCTCTG	GAATAGTGAA	CTGGGACATG	GGGGAAAGTCA	GTCTCTCTCT	1260
GCCACAAGCCA	CAGATTAAAA	ATAATAATGT	TAACTGATCC	CTAGGCTAAA	ATAATAGTGT	1320
TAACTGATCC	CTAAAGCTAA	AAAGTTCTTT	TGTTAATTCA	GGTGATGGCA	GCAAGACCCA	1380
TCTTAAGGAT	AGACTAAGTT	TGCTTAGTTT	GAGGTCATAT	CTGTTTGCTC	TCAAGCATGT	1440
ACTGGAAGAA	GTTCATCAC	ACAGCTCTCA	GGACTGCTCT	CCTCTCTACA	GCAATGGATA	1500
ATGCTTCACT	AGCCTTTTGA	GATAATTTTT	GATCAGAGAA	AAAACCTTGA	GCTGGGCCAA	1560
AAAGGAGGAG	CTTCAACCTG	TGTGCAAAAT	CTGGGAACCT	GACAATATAG	GTGGGGGGCC	1620
AGGATGAGGA	AAAAGGAAGC	GGAAAGACCT	GCCACCCCTT	CTGTTAAGGA	GGCCCCGTGA	1680
TCAAGCTCCA	CCATTTGCA	TGCTGGCTAT	CCCAAGAGCT	TACATAAAGG	GACAATTGGA	1740
GCCTGAGAGG	TGACAOTGCT	GACACTACAA	GGCTGGGAGC	TCCGGGCACT	CAGACATC	1798
ATG AGT TGG TCC TTG CAC CCC CGG AAT TTA ATT CTC TAC TTC TAT GCT	Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala	1846				
1 5 10 15						
CTT TTA TTT CTC TCT TCA ACA TGT GTA GCA GTAAAGTGTG TCTTCACAAA	Leu Leu Phe Leu Ser Ser Thr Cys Val Ala	1896				
20 25						
ACGTTGTTTTA AAATGGAAAG CTGAAAAATA AAACAGATAA TAACTAAGTG AAATTTTCGT		1956				
ATTTTTTCTC TTTTAA TAT GTT GCT ACC AGA GAC AAC TGC TGC ATC TTA		2005				
Tyr Val Ala Thr Arg Asp Asn Cys Cys Ile Leu	30 35					
GAT GAA AGA TTC GTAAGTAATT TTTATGTTTC TCCCTTTGTG TGTGAACCTGG		2057				
Asp Glu Arg Phe						

-continued

40

AGAGGGGOCAG	AGGAATAGAA	ATAATTCCCT	CATAAATATC	ATCTGGCACT	TOTAACITTT	2117
TAAAAACATA	GTCTAGGTTT	TACCTATTTT	TCTTAATAOA	TTTTAAGAOT	AGCATCTOTC	2177
TACATTTTTA	ATCACTOTTA	TATTTTCAO	OOT	AOT	TAT TOT CCA ACT ACC TOT	2230
			Gly	Ser	Tyr Cys Pro Thr Thr Cys	45
GOC ATT GCA GAT TTC CTG TCT ACT TAT CAA ACC AAA GTA GAC AAG GAT	2278					
Gly Ile Ala Asp Phe Leu Ser Thr Tyr Glu Thr Lys Val Asp Lys Asp						
50 55 60 65						
CTA CAG TCT TTG GAA GAC ATC TTA CAT CAA GTT GAA AAC AAA ACA TCA	2326					
Leu Glu Ser Leu Glu Asp Ile Leu His Glu Val Glu Asn Lys Thr Ser						
70 75 80						
GAA GTC AAA CAG CTG ATA AAA GCA ATC CAA CTC ACT TAT AAT CCT GAT	2374					
Glu Val Lys Glu Leu Ile Lys Ala Ile Glu Leu Thr Tyr Asn Pro Asp						
85 90 95						
GAA TCA TCA AAA CCA A GTGAGAAAAT AAAGACTACT GACCAAAAAA	2420					
Glu Ser Ser Lys Pro						
100						
TAATAATAAT AATCTGTGAA GTTCTTTTGC TOTTOTTTTA GTTGTICTAT TTGCTTAAAG	2480					
ATTTTTATGT CTCTGATCCT ATATTACAG AT ATG ATA GAC OCT OCT ACT TTO	2532					
			Asn Met Ile Asp Ala Ala Thr Leu			
			105 110			
AAO TCC AGO ATA ATG TTA GAA GAA ATT ATG AAA TAT GAA GCA TCG ATT	2580					
Lys Ser Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile						
115 120 125						
TIA ACA CAT GAC TCA AGT ATT CG GTAAAGGATTT TTGTTTTAAT TTGCTCTOCA	2633					
Leu Thr His Asp Ser Ser Ile Arg						
130						
AGACTGATTT AGTTTTTATT TAATATTCTA TACTTGAOTG AAAGTAAATTT TTAATGTGTT	2693					
TTCCCCATTT ATAATATCCC AGTGACATTA TGCCTGATTA TOTTGAGCAT AGTAGAGATA	2753					
GAAGTTTTTA GTGCAATATA AATTATACTG GOTTATAATT GCTTATTAAAT AATCACATTO	2813					
AAGAAAAGATG TTCTAGATGT CTTCAAATGC TAGTTTGACC ATATTTATCA AAAATTTTTT	2873					
CCCCATCCCC CATTTATCTT ACAACATAAA ATCAATCTCA TAGGAAATTTG GGTGTTGAAA	2933					
ATAAAATCCT CTTTATAAAA ATGCTGACAA ATTGOTGOTT AAAAAAATTA GCAAGCAAGAG	2993					
GCATAGTAAO GATTTTGGCT CCTAAAGTAA ATTATATTGA ATGTGGAGCA GGAAGAAACA	3053					
TGTCTTGAOA GACTAAGTOT GGCAAAATAT GCAAAAGCTCA TATTGATCAT TGCAGAAATOA	3113					
ACCTGCATAG TCTCTTCCCT TCATTTGAAA GTGAATGTCT CTGTTAAAGC TTCTCAAGGA	3173					
CTCATAAACT TTCTGAACAT AAGGTCTCAG ATACAGTTTT AATATTTTTT CCAATTTTTT	3233					
TTTTCTGAAT TTTTCTCAAA GCAAGCTTGAO AAATTOAGAT AAATAOTAOC TAGGAGAGAG	3293					
TGCCCCAAGA AAGATTTCTC CTCTTTTTGC TATCAGAGGG CCCTTOTTAT TATTGTTATT	3353					
ATTATTACTT GCATTATTAT TOTCCATCAT TGAAGTTGAA GGAAGTTATT GTACAGAAAT	3413					
TGCCTAAOAC AAGGTAGAGG GAAAAACGTGG ACAAATAOTT TGTCTACCCT TTTTACTTC	3473					
AAAGAAAAGAA CGTTTTATGC ATTGTAGACA GTTTTCTATC ATTTTTGGAAT ATTTGCAAGC	3533					
CACCCGTOTAA GTAACATCAA AAGGAGGGTT TTTACTTCCC CCAATCCATT CCCAAAAGCTA	3593					
TOTAACCAOA AGCATTAAAA AAGAAAAGGGG AAGTATCTGT TTTTTTATTT TACATACAAT	3653					
AACGTTCCAG ATCATGTCCC TGTGTAAOTT ATATTTTAGA TTGAAAGCTTA TATGTATAOC	3713					
CTCAOTAGAT CCACAAOTGA AAGGTATACT CCTTCAAGCAC ATGTGAATTA CTGAAGTGAO	3773					

GCTTGGGCTG

CTTTTCCTGC	TTCTAAAGCA	TCAGGGGGGTG	TTCCTATTAA	CCAGTCTCOC	CACTCTTCOA	3833
GGTTOGTATC	TGCTGTCCCT	TATGCATAAA	GTAAAAAGCA	AAATGTCAAT	GACATTTGCT	3893
TATTOGACAA	GACTTTGTTA	TTTGTGTTGG	GAGTTGAGAC	AATATGCCCC	ATTCTAAGTA	3953
AAAAAGATTCA	GGTCCACATT	GTATTCCTGT	TTTAATTGAT	TTTTTGATTT	GTTTTTCTTT	4013
TTCAAAAAAGT	TTATAATTTT	AATTCATOTT	AATTTAGTAA	TATAATTTTA	CATTTTCCTC	4073
AAAGAAAGAA	TAATTTATCA	GAAAGCACTT	CTTAAGAAAA	TACTTAAGCA	TTTCCAAAAA	4133
AAATATAAAA	TTACTCTTCT	GAAAGGAATA	CTTATTTTTG	TCTTCTTATT	TTTGTATCT	4193
TATGTTTCTG	TTTGTAG A	TAT TTG CAG	GAA ATA TAT	AAT TCA AAT	AAT CAA	4244
		Tyr Leu Glu Glu Ile	Tyr Asa Ser Asa Asa Glu			
		135	140		145	
AAO ATT GTT	AAC CTG	AAA GAG AAG	GTA GCC CAG	CTT GAA GCA	CAG TGC	4292
Lys Ile Val	Asa Leu	Lys Glu Lys Val	Ala Glu Leu Glu Ala Glu Cys			
	150		155		160	
CAG GAA CCT	TGC AAA	GAC ACO GTG	CAA ATC CAT	GAT ATC ACT	GGG AAA G	4341
Glu Glu Pro	Cys Lys Asp Thr	Val Glu Ile His	Asp Ile Thr Gly Lys			
	165	170	175			
GTAACTGATG	AAGGTTATAT	TGGGATTAGG	TTCATCAAAO	TAAOTAATOT	AAAAGGAGAAA	4401
GTATGTACTG	GAAAAGTATAG	GAATAGTTTA	GAAAAGTGGCT	ACCCATTAAO	TCTAAGAAAT	4461
TCAOTTOTCT	AGACCTTTCT	TGAATAGCTA	AAAAAAAACAG	TTTAAAAAGGA	ATGCTGATOT	4521
GAAAAAGTAAO	AAAATTATTC	TTGGAAAAATG	AATAGTTTAC	TACATGTTAA	AAOCTATTTT	4581
TCAAGGCTGG	CACAGTCTTA	CCTGCATTTT	AAACCACAGT	AAAAOTCGAT	TCTCCTTCTC	4641
TAG AT TOT	CAA GAC ATT	GCC AAT AAO	GGA OCT AAA	CAG AGC GGG	CTT	4688
Asp Cys	Glu Asp Ile Ala Asa	Lys Gly Ala	Lys Glu Ser Gly Leu			
	180	185	190			
TAC TTT ATT	AAA CCT CTG	AAA OCT AAC	CAG CAA TTC	TTA GTC TAC	TOT	4736
Tyr Phe Ile	Lys Pro Leu Lys	Ala Asa Glu Glu	Phe Leu Val Tyr Cys			
	195	200	205			
GAA ATC GAT	GGG TCT GGA	AAT GGA TGG	ACT GTG TTT	CAG AAG		4778
Glu Ile Asp	Gly Ser Gly Asa Gly	Trp Thr Val	Phe Glu Lys			
	210	215	220			
GTAAATTTTTT	CCCCACCATG	TGTATTTAAT	AAATTCCTAC	ATTGTTTCTG	CCATATGACA	4838
GATACTTTTC	TAAGCACCTT	GTGAACCOTA	GCTCATTTAA	TCCTTTOCAAT	AGCCCTAAGG	4898
GGAAAGTACT	TCTGTTACTC	CTATTTACAG	AAAAAGAAAC	TGAGGCACAC	AAAGTTAAAT	4958
AACTTCCCCA	AGACCACATA	ACTAATAAGC	AACAGAGTCA	GCATTTGAAC	CTAGGCAOTA	5018
TAGTTTCAOA	GTTTGTGACT	TGACTCTATA	TTGTACTGGC	ACTGACTTTG	TAGATTTCAT	5078
GTGGCACATA	ATCATAGTAC	CACAGTGACA	AATAAAAAAG	AGGAAACTCT	TTTGTCAAGT	5138
AGGTCAAAGAC	CTGAGGTTTC	CCATCACAAO	ATGAGGAAAG	CCAACACCAC	CCCCCACCAC	5198
CCCCACCACCA	TCACCACCCT	TTCACACACC	AGAGGATACA	CTTGGGCTGC	TCCAAGACAA	5258
GGAACTGTG	TTGCATCTGC	CACTTGCTGA	TACCCACTAG	GAATCTTGGC	TCCTTTACTT	5318
TCTGTTTACC	TCCCACCACCT	GTATATAACTG	TTTCTACAGG	GGGCGCTCAG	AGGGAATGAA	5378
TGGTGGAAAGC	ATTAGTTGCC	AGACACCGAT	TGAGCAATGG	GTTCCATCAT	AAGTGTAAGG	5438
ATCAOTAATA	TCCAAGTAGA	GTTCGTGAAGT	CGTCTAGGTO	TCTTTTTAAT	ATTACCACTC	5498
ATTTAGAAAT	TATGATGTGC	CAGAAACCCT	CTTAAGTATT	TCTCTTATAT	TCTCTCTCAT	5558
GATCCTTCCA	GCAACCCTAA	GAAOTAACCA	TCATTTTTTC	TATTTGATAC	ATGAGGAAAC	5618
TGAGGTAGCT	TGGCCAAAGAT	CACTTAGTTG	GGAGTTGATA	GAACCAAGTGC	TCTGTATTTT	5678
TGACAAAAATG	TTGACAAGCAT	TCTCTTTACA	TGCATTGATA	GTCTATTTTC	TCCTTTTGTCT	5738

-continued

CTTCCAAATG TGTAAATTAG AGA CTT GAT GGC AGT GTA GAT TTC AAG AAA AAC	5790
Arg Leu Asp Gly Ser Val Asp Phe Lys Lys Asn	
225 230	
TGG ATT CAA TAT AAA GAA GGA TTT GGA CAT CTG TCT CCT ACT GGC ACA	5838
Trp Ile Glu Tyr Lys Glu Gly Phe Gly His Leu Ser Pro Thr Gly Thr	
235 240 245	
ACA GAA TTT TGG CTG GGA AAT GAG AAG ATT CAT TTG ATA AGC ACA CAG	5886
Thr Glu Phe Trp Leu Gly Asn Glu Lys Ile His Leu Ile Ser Thr Glu	
250 255 260 265	
TCT GGC ATC CCA TAT GCA TTA AGA GTG GAA CTG GAA GAC TGG AAT GGC	5934
Ser Ala Ile Pro Tyr Ala Leu Arg Val Glu Leu Glu Asp Trp Asn Gly	
270 275 280	
AGA ACC AG GTACTGTTTT GAAATGACTT CCAACTTTTT ATTGTAAAGA	5982
Arg Thr Ser	
TTGCCTGGAA TGTGCACTTT CCAACTATCA ATAGACAATG GCAAATGCAO CCTGACAAAT	6042
GCAAACAGCA CATCCAAGCA CCATTTTCTC CAGGAGTCTG TTTGTTTCTT GGGCAATCCA	6102
AAAAAGTAAA TTCTATTCAO GATGAATCTA AGTGTATTGG TACAATCTAA TTACCCTGGA	6162
ACCATTGAGA GTAATAAGTA ATTACTGAAC TTTTAATCAO TCCCAAGGAAT TGAGCATAAA	6222
ATTATAATTT TATCTAGTCT AAATTACTAT TTCAAGAAAG AGGTATTATT ATTAATCCCA	6282
TTTTATAOAT TAACTTGCTC AAAATCACAT TGCTGATAAG TGGTAGAGGT AGAATTCAOA	6342
CTCAAAGTAT TTAACTTTAA AGCCTGTCTT CTTAACAACAT ATCCTGTTTG AAAAGCAAAAT	6402
ACAAGCTCTT CAGACTTCTC AGTGCCTTGA TGGCCATTTA TTCTGTCAAA TCATGAAGTA	6462
CCCTAAAAAT AAACCAAGTA GCTCTTTTGA TGATCTAGAG GCTTCTTTTT GCTTGAAGTA	6522
TTTGAAGGTT TTAAGCATTO TTACCTAATT AAAATGCAOA AAAATATCCA ACCCTCTTGT	6582
TATGTTTTAA GAAATGTOAA ATATATTGTC TTCAAAACACA TGGACTTTTT TTTATTGCTT	6642
GTTTGGTTTT TAATCCAGAA AGTGCTATAO TCAATAGACC TTCTTCTAAG AAAAGACCTT	6702
CCATTTCCCA GGCAGTGAAG ATTAGAAAAAT AAGCTAAATA TTTTCTGGAAT ATTTCTGTTT	6762
ATTGATTAAO GGCATCTCTT TCCCCACTC TATAGAAATG TTGTCCACTT GCACAATTTT	6822
TTCCAGGAAA GAATCTCTCT AACTCCTTCA GCTCACATGC TTTGAGCCAC ACAGGGAAGA	6882
CTTTGATTGT GTAAATGCCCT CAGAAAGCTCT CTTTCTTGCC ACTACCACAC TGATTTGAGG	6942
AAGAAAAATCC CTTTAGCACC TAACCCTTCA GTTGCTATGA GTGGCTAATG GAAGTGTACC	7002
TCCTTCAAGT TTTGTGCAAT AATTAAAGGT CACTCACTGT CAGATACTTT CTGTGATCTA	7062
TGATAATGTO TGTGCAACAC ATAACATTTT AATAAAAAAT GAAAAATATGA AATTAGAGTC	7122
ATCTACACAT CTGGAATTTGA TCTTAGAATG AAACAAAGCA AAAAAATATC AAGTGAAGTC	7182
AATTATTAGT TTTCAAGGAT GCTTCAAAAG CTTCTAGGCC CATCCCGGGA AGTGTTAATG	7242
AGCTGTGAGC TGTTTCACAT ATCTATTGCC TCTTGCCAGA TTTGCAAAAA ACTTCACTCA	7302
ATGAGCAAAAT TTCAAGCTTA AAAAAACAAA TCAAAAAATC CAAAGGAGCA TCCTACGAAA	7362
GAAGGAAACTT CTGAGATCCC TGAGGAGGAT CAGCATGTGA TGGTTGTATT TCCTTCTTCT	7422
CAG T ACT GCA GAC TAT OCC ATG TTC AAG GTG GGA CCT GAA GCT GAC	7468
Thr Ala Asp Tyr Ala Met Phe Lys Val Gly Pro Glu Ala Asp	
285 290 295	
AAG TAC CAC CTA ACA TAT OCC TAC TTC GCT GGT GGG GAT GCT GGA GAT	7516
Lys Tyr Arg Leu Thr Tyr Ala Tyr Phe Ala Gly Gly Asp Ala Gly Asp	
300 305 310	
OCC TTT GAT GGC TTT GAT TTT GGC GAT GAT CCT AGT GAC AAG TTT TTC	7564
Ala Phe Asp Gly Phe Asp Phe Gly Asp Asp Pro Ser Asp Lys Phe Phe	
315 320 325 330	

66570-3673260

-continued

ACA	TCC	CAT	AAT	GGC	ATG	CAG	TTC	AGT	ACC	TGG	GAC	AAT	GAC	AAT	GAT	7612
Thr	Ser	His	Asn	Gly	Met	Gln	Phe	Ser	Thr	Trp	Asp	Asn	Asp	Asn	Asp	
				335					340					345		
AAO	TTT	GAA	GGC	AAC	TGT	GCT	GAA	CAO	GAT	GGA	TCT	GGT	TGG	TGG	ATG	7660
Lys	Phe	Gln	Gly	Asn	Cys	Ala	Gln	Gln	Asp	Gly	Ser	Gly	Trp	Trp	Met	
			350					355					360			
AAC	AAO	TGT	CAC	GCT	GGC	CAT	CTC	AAT	GGA	GTT	TAT	TAC	CAA	G		7703
Asn	Lys	Cys	His	Ala	Gly	His	Leu	Asn	Gly	Val	Tyr	Tyr	Gln			
		365					370				375					
GTATGTTTTT	CTTTCCTTAA	TTCCAAATTA	ATGTATAGTG	TATACTATTT	TCATAAAAAA											7763
TAATAAATAG	ATATGAAAGAA	ATGAAAGATA	ATTTATAAAO	ATAATAAGGA	TTTTATCATG											7823
TTCTTTATTT	CAACTAAATT	CTTTOAAACT	GGAAATGGAAT	AATACCAAGT	TCATGCTTAA											7883
AATTAGCCCT	TCTAAAAGAAA	TCCACCTGCT	GCAAAATATC	CAATAATTTG	GCATTATATG											7943
TGAAACTATC	ACCATCATAG	CTGGCACTGT	GGTTGTGGG	ATCTCCTTTA	GACATACAAC											8003
ATAAATGATC	TGGATGGAAT	AACATTACTA	CATGGAATGCT	TGTTGACACA	TAACTCTGGC											8063
TTCCCATGAG	CTTTGTGTCA	GATACACGCA	GTGAAACAGT	GTTTGGAAGG	ACAGAATAAA											8123
GAGAAAGCAA	GCACTGGAAT	GGGCAAGGGT	TTGTGAAAGC	TTGAGAGAAO	AGACCAATCT											8183
GAGGACAATA	GACACTTATT	TTAGGATGGA	GGTTGGATGA	GGAGGCTATA	GTTCCTTATA											8243
AGCTTGGAAAT	GGTTTGGAAAC	ACTGTTTTCA	CTCACCTACC	CAAGCAATTAT	GTGTGGGGAA											8303
GCCTTACCGA	TGCTAAAAGG	TCCATGTTAC	AATAATGGA	TTATTTGGA	ATCCCAATGG											8363
TATTCATGA	ATAAAACCAC	TATGAAAGTA	ATCCCACTCA	ACAAGACTCTC	COTGGAGAAA											8423
GGACAGCAAC	ACCACCTTGG	GAAAAGCCAAA	CAATCAAGACC	AGACCTGTTT	AGCATCAATA											8483
GGACTTCCCT	ACCATATCTG	CTGGGTAGAT	GAGTGAAGACC	AGTGTICCAA	ACCACCTCCG											8543
GCCTTGTAGCA	AACCATAATC	TCCTCATCTA	CCAAAGATGA	CAACCTTACC	TCCTGATGTC											8603
CTAGCCAATC	ACCAACTAAG	AAACTTTTGA	CAATTTATTT	AAAATAACAG	TTTGATTTTC											8663
ACAATATTTT	TAAATTTGAG	AAACATAACT	TATCTTTTGA	CTCACAAACC	ACATAATGAG											8723
AAGAAACTCT	AAGGAAAAAT	GCCTTATCTG	TGTGACCCCG	GGGCGCATGC	CAGAGCTGTA											8783
GTTCATGCCA	GTGTTGTGCT	CTGACAAAGC	TTTTACAGAA	TTACATGAG	TCTGCTTCCC											8843
TAGGACAAAG	AGAAAGCCAAA	TCAACAGAGG	CTGCACTTTA	AAATGGAAGC	ATAAAATAAC											8903
ATGCCAGAAC	CATTTCTTAA	AGCTCCTCAA	TCAACCAACA	AAATTTGTGCT	TTCAAATAAC											8963
CTGAGTTGAC	CTCATCAAGG	ATTTGTGTC	TCCTTCTCTT	CTAACCTGCC	TGAAAGAAAG											9023
TGTTCCACAG	CAAGTGAATC	CGGATGGAAT	AACTTTAGGG	ACAAGAGGCCA	ATTAGGGAAC											9083
TTTGAGTTTC	TAGCCCTACT	AGTATGGAAT	AAATTTAAAA	TGTGATGTC	ACTATGATC											9143
ACAGCACAGG	TGTTGTTTAA	TAAATGTTT	ATTTATATA	TTGATATTTT	AGGAATCTTT											9203
GGAGATATTT	TCAATTAAGC	GATAATACTA	TAAATTTTAT	GTAACTGGCA	ATGCACTTCC											9263
TAAATGAGAG	CTCTTCATAG	ACTTGAGAG	GTAAAAAGAT	TCCAAGATAA	TGATATGTAC											9323
ATCTAGCACT	TGTTTTAG	GT GGC ACT	TAC TCA AAA	GCA TCT ACT	CCT AAT											9373
		Gly Gly Thr	Tyr Ser Lys	Ala Ser Thr	Pro Asn											
			380		385											
GGT TAT GAT AAT GGC ATT ATT TGG OCC ACT TGG AAA ACC CGG TGG TAT																9421
Gly Tyr Asp Asn Gly Ile Ile Trp Ala Thr Trp Lys Thr Arg Trp Tyr																
	390						395				400					
TCC ATG AAG AAA ACC ACT ATG AAG ATA ATC CCA TTC AAC AGA CTC ACA																9469
Ser Met Lys Lys Thr Thr Met Lys Ile Ile Pro Phe Asn Arg Leu Thr																
	405					410					415					
ATT GGA GAA GGA CAG CAA CAC CAC CTG GGG GGA GCC AAA CAG GTC AGA																9517

-continued

Ile Gly Glu Gly Glu Glu His His Leu Gly Gly Ala Lys Glu Val Arg
 420 425 430 435
 CCA GAG CAC CCT GCG GAA ACA GAA TAT GAC TCA CTT TAC CCT GAG GAT 9565
 Pro Glu His Pro Ala Glu Thr Glu Tyr Asp Ser Leu Tyr Pro Glu Asp
 440 445 450
 GAT TTT TAGAAAATTA ACTGCTAACT TCTATTOACC CACAAAATTT CAGAAAATTCT 9621
 Asp Leu
 CTGAAAATTT CTTCCTTTTT TCTCTTACTA TATTTATTGA TTTCAAGTCT TCTATTAAAG 9681
 ACATTTAGCC TTCAATGGAA ATTAAAACTC ATTTAGGACT GTATTTCCAA ATTACTGATA 9741
 TCAGAGTTAT TTA AAAAATTO TTTATTTGAG GAGATAACAT TTCAACTTTG TTCCTAAATA 9801
 TATAATAATA AAATGATTGA CTTTATTTGC ATTTTATGA CCACCTGTCA TTTATTTTGT 9861
 CTTCTAAAT TATTTTCATT ATATCAATA TTTAGTATG TACTTAATAA AATAGGAGAA 9921
 CATTTTAGAG TTTCAAATTC CCAGGTATTT TCCTTGTTTA TTACCCCTAA ATCATTCTTA 9981
 TTTAATTCIT CTTTTTAAAT GGAGAAAATT ATGTCTTTTT AATATGGTTT TTGTTTTGTT 10041
 ATATATTAC AGGCTGGAGA CTTTTAAAA ACCCTTTCAA AAGAGATTTA CTTTTTTAAA 10101
 GGACTTTATC TGAACAGAGA GATATAATAT TTTTCCTATT GGACAATGGA CTTCGAAAAGC 10161
 TTCATTTCAT TTTAAGAGCA AAAGACCCCA TTTGAAAAAC TCCATAACAG TTTTATGCTG 10221
 ATGATAATTT ATCTACATGC ATTTCAATAA ACCTTTTGT TCTAAGACT AGATACATGG 10281
 TACCTTTATT GACCATTAAA AAACCACCAC TTTTGGCCAA TTTACCAATT ACAATTGGGC 10341
 AACCATCAOT AOTAAITGAG TCCTCATTTT ATGCTAAATG TTATGCCTAA CTCTTTGGGA 10401
 GTTACAAAAG AAATAOCAAT TATGGCTTT OCCCTCTAGG AGATACAGGA CAAATACAGG 10461
 AAAATACAGC AACCCAAACT GACAATACTC TATACAAAGAA CATAATCACT AAGCAGGAOT 10521
 CACAGCCACA CAACCAAGAT GCATAGTATC CAAAGTGCAO CTG 10564

(2) INFORMATION FOR SEQ ID NO.:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO.:

Met Ser Trp Ser Leu His Pro Arg Asn Leu Ile Leu Tyr Phe Tyr Ala
 1 5 10 15
 Leu Leu Phe Leu Ser Ser Thr Cys Val Ala Tyr Val Ala Thr Arg Asp
 20 25 30
 Asn Cys Cys Ile Leu Asp Glu Arg Phe Gly Ser Tyr Cys Pro Thr Thr
 35 40 45
 Cys Gly Ile Ala Asp Phe Leu Ser Thr Tyr Glu Thr Lys Val Asp Lys
 50 55 60
 Asp Leu Glu Ser Leu Glu Asp Ile Leu His Glu Val Glu Asn Lys Thr
 65 70 75 80
 Ser Glu Val Lys Glu Leu Ile Lys Ala Ile Glu Leu Thr Tyr Asn Pro
 85 90 95
 Asp Glu Ser Ser Lys Pro Asn Met Ile Asp Ala Ala Thr Leu Lys Ser
 100 105 110
 Arg Ile Met Leu Glu Glu Ile Met Lys Tyr Glu Ala Ser Ile Leu Thr
 115 120 125
 His Asp Ser Ser Ile Arg Tyr Leu Glu Glu Ile Tyr Asn Ser Asn Asn
 130 135 140

GATTTAGCC

Table 1

Variable	Mean	SD	Range
Age	67.8	9.0	45-85
Gender	Male		
Females	10		
Males	10		
Marital status	Married		
Single	10		
Divorced	10		
Widowed	10		
Educational level	High school		
Bachelor's degree	10		
Master's degree	10		
PhD	10		
Occupation	Retired		
Employed	10		
Unemployed	10		
Health status	Good		
Fair	10		
Poor	10		
Functional status	Independent		
Dependent	10		
Cognitive function	Normal		
Impaired	10		
Quality of life	High		
Low	10		

ACCGGTGTCTG	ACCTGCAAGT	CAACGGAATCT	CTGTGTCTGT	TTTCATGTTA	GTACCACACT	60
GTTTTGGTGG	CTGTAGCTTT	CAGCTACAAT	CTGAAGTCAT	AAAAGCCTGT	ACCTCCAAGT	120
CTGTTCTCTC	TCAAGATTGT	GTTCTGCTGT	TGGGTCITT	AGTGTCTCCA	CACAATTTT	180

-continued

AGAAATTOTTT	OTTCTAOTTC	TGTAAAAAAT	GATGCTGOTA	TTTTGATAAG	GATTGCATTO	240
AATCTOTAAA	GCTACAGATA	TAOTCATTGG	GTAOTACAGT	CACITTAACA	ATATTAACCTC	300
TTACACATCTG	TGAGCATGAT	ATATTTTCCC	CCTCTATATC	ATCTTCAATT	CCTCCTATCA	360
GTITCTTTCA	TTGCAOTTTT	CTGAGTACAG	GTCTTACACC	TCCTTGGTTA	GAGTCATTCC	420
TCAGTATTTT	ATTCCITTTG	TACAATTGTO	AATGAGGTAA	TTTTCTTAOT	TTCTCTTTCT	480
GATAOCTCAT	TOTTAAGTGA	TATATAAAAA	AGCAACAAGT	TTCTATGTAT	TAATTTTGTG	540
TCCTGCAACA	GATTTCTATG	TATTAATTTT	GTATCCTGCT	ACTTTACGGA	ATTCACCTAT	600
TAGCTTTTTG	GTGACATCTT	GAGGATTTTC	TGAAAAAAAT	GGCATGGTAT	GGTAGGACAA	660
GGTGTGATGT	CATCTGCAAA	CAGTGGCAGT	TTTCCTTTCT	CCCTTCCAAC	CTGGATTTC	720
TTGATTTCTT	TCTGTCTGAG	TACGACTAAG	ATTCCCAATA	CTATACCGAA	TAAAAAGTGG	780
AAGAGTGGAC	ATCCTTGTCT	TATTTTTCTG	ACCTTAGAGG	AAATGCTTTC	AGTTTTTCAC	840
CATTAAATTAT	AATGTTTACT	GTGGGCTTGT	CATATGTGGC	CTTCATTATA	TGGAAGTCTA	900
TTCCCTCTAT	ACCCACCTTG	TTGAGAGTTT	TTATCATAAA	AGTATGTTGA	ATTTTGTCAA	960
AAGTTTTTTC	TGCATCTATT	GAGATGATTT	TTACTCTTCA	ATTCATTAAT	GATTTTTATT	1020
CTTCATTTTG	TAAATGATTT	CCATTCTTCA	ATTTGTTAAC	GTGGTATATC	ACATTGATTO	1080
ATTTGTGGAT	ACCTTTGTAT	CCCTGGGATA	AACCTCACTT	GATCATGAGC	TTTCAATGTA	1140
TTTTTGAATT	CACTTTGTCT	ATATTCTGTT	GGGTATTTTT	GCATCTCTAT	TCATCAATGA	1200
TATTTGGCTA	AGAAAAGTTT	TGTCTGTTTT	TAGTATCAAG	GTGATGCTGG	CCTCATAGAG	1260
AGAGTTTAGA	AGCATTTCCT	CCTCTTTGAT	TTTTCGGAAT	AGTTTGAOTA	GGATAGGTAT	1320
TAACTCTTCT	TAAATGTTTT	GGGGACTTCC	CTGGTGAAGC	GGTGGTTGAG	AATCCGCTTC	1380
AAGGATGTGG	GTTTGATCCC	TGGTCAAGGA	ACCATTAATA	AGATCCACAA	TGCTGCAAGC	1440
AACAAAGCCC	CAGGCTGCAA	CCACTGAGCT	GCAACCGCTG	CAGTGGCCAC	AGGCCACGAC	1500
CAGAGAAAGC	CCACATACAG	CAGGGAAGAC	CCAACACAAC	CGGAAAAAGG	AGTTTGGTGG	1560
AATACAGCTG	TGAAAGCCGC	TGGTCCCTGA	CTCCTGCTTG	AGGGAATTTT	TAAAAAATTA	1620
TTGATTCAAT	TTCACTACTG	GTAACTGGTC	TGTTCAATAT	TTCTATTTCT	TCCGGGTTCA	1680
GTCTTGGGAG	ATTGTACATG	CCTAGGAATG	TGTCCGTTTC	TTCTAGGTTG	TCCATTTTAT	1740
TGGACATGCA	TGGGAAGACA	CAGCACCAGC	CAGCGAGACT	CATGCTGGCT	TCCTGGGGCC	1800
AAGCTGGGGC	CCCAAGCAAG	ATGGCATCCT	AGAATGTGTG	AAAAGCCACT	GACCCCTGCCC	1860
AGCCCCACAA	TTTCATTCTG	AGAAATGATT	CCTTGGCTTCT	GCACTTACAG	GCCCCAGGATC	1920
TGACCTGCTT	CTGAGGAAGC	GGGGTTTTTG	CAGGACGGGG	AGATGCTGAG	AGCCGACGGG	1980
GGTCCAGGTC	CCCTCCCAAG	CCCCCCTGTC	TGGGGCAGCC	CTTGGGAAAG	ATTGCCCCAG	2040
TCTCCCTCCT	ACAATGGTCA	GTCCCAAGCTG	CCCCAGGCCA	GAGCTGCTTT	ATTTCCGTCT	2100
CTCTCTCTGG	ATGGTATTTCT	CTGGAAGCTG	AAAGTTTCCTG	AAATTATGAA	TAGCTTTGCC	2160
CTGAAAGGCA	TGGTTTGTGG	TCACGGTTCA	CAGGAACCTG	GGAGACCCTG	CAGCTCAAGC	2220
GTCCCGAGAT	TGGTGGCACC	CAGATTTCCCT	AAGCTGCTG	GGGAACAAGG	CGCTTGTTC	2280
TCCCTGGCTG	ACCTCCCTCC	TCCCTGCATC	ACCAAGTTCT	GAAAGCAAGG	CGGTGCTGGG	2340
GTCAAGCCCT	CTGCACTCTA	ACGCCGGTGT	CCAAACCACC	CGTGGTGGTG	TTGGGGGGGC	2400
TACCTATGGG	GAAAGGCTTC	TCACTGCAAT	GGTGGCCCCC	GTCCCTCTG	AGATCAAGAG	2460
TCCCAAGTCCG	GACGTCAAAAC	AGGCCGAGCT	CCCTCCAGAG	GCTCCAGGGA	GGGATCCTTG	2520
CCCCCCCCCT	GCTGCTTCCA	GCTCCTGGTG	CCGACCCCTT	GAGCCTGATC	TTGTAGACGC	2580

CTCAGTCTAG	TCTCTGCCTC	CGTGTTCACA	CGCCTTCTCC	CCATGTCCCC	TCCGTGTCCC	2640
COTTTTCTCT	CACAAAGGACA	CCGAGACATTA	GATTAAGCCCC	TOTTTCCAGCC	TCACCTGAAAC	2700
AGCTCACATC	TGTAAAGGACC	TAGATTCCAA	ACAAAGATTCC	AACCTGAAAT	TCCCAGGTGGA	2760
TGTGAGTTCT	GCGCGGACAT	CCTTCAACCC	CATCACAGCT	TGCAOTTCAI	CGCAAAACAT	2820
GAACCTGCG	GTTTATCGTA	AAACCCAGGT	TCTTCATGAA	ACACTGAGCT	TCGAGGCTTG	2880
TTGCAAGAAAT	TAAAGGTGCT	AATACAGATC	AGGCGAAAGGA	CTGAAAGCTG	CTAAAGCCTCC	2940
TCTTTCCATC	ACAAGGAAAGG	GCGCGCTGCG	GCGCGCTGGA	GCTCTGCTCC	CGTGAAGTGA	3000
CTCTTTCTTG	CTACAGTCAC	CAACAGTCTC	TCTGCGAAAG	AAACCAAGAG	CCAGAGAGCA	3060
AGCCGGAAGCT	AGTTTAAAGAG	ACCCCTGAAC	CTCCACCCAA	GATGCTGACC	AGCCAAGCGGG	3120
CCCCCTGAAA	AGACCTTACA	GTTCAAGGGGG	GAAAGAGGGGG	TGACCCGCGCA	GCTCCCTGCT	3180
ATCAGGAGAC	ATCCCCGCTA	TCAAGAGATT	CCCCACCTT	GCTCCCGTTT	CCCTATCCCA	3240
ATACGCCCCAC	CCCACCCCTG	TGATGAGCA	TTTAAOTCACT	TAGAAOTGCA	ACTGAAAGGCT	3300
TTTGCATCCC	CTTTGCGAGA	GCGACAAAGG	ACCCACAGCC	TGCTGCGTAC	CGACGCGCAT	3360
GTGGAITCAG	CCAAGAGGGC	TGTCTGTCAC	CCTCCCTGCT	CGCGCCCCCT	CTGTGCTCAG	3420
CAACACACCC	AGCAGCAAGCA	TTCCCGCTGC	TCCTGAGGTC	TGCAAGCAGC	TCGCTGTAGC	3480
CTGAGCGGTG	TGAGAGGAAAG	TGTCTGAGGA	GATTTAAAAAT	GTGAGAGGGG	GGAAGGTGAGGA	3540
GOTTGCGCCC	TGTGCGGCTG	CCCATCCCAAC	GTGCTGTCAT	TAGCCCCAGT	GCTGCTCAGC	3600
CGTGCCCCCG	CGCGAGGGGT	CAGGTCACTT	TCCCGTCTTG	GCGTTATTTAT	GACTCTTGTG	3660
ATTGCCATTG	CCATTTTTTG	TACCTTAACT	GCGCAAGCAGG	TGCTTGAGGA	GCCCTCGATA	3720
CGGACCAAGGT	CCTCCCTCGG	AGCTCGACCT	GAACCCCATG	TCACCCCTTG	CCCAGCCTTG	3780
AGAGGGTGCG	TGACTGAGGA	GATCCCTTCA	CCCAAGGGCA	CGGTACAGTG	GTTTGGAAGGA	3840
GCTGCTGCCC	AAAGCAGAGG	CCACCCCTCA	GACACACCTT	GTCCCCAGTG	CTGGCTCTGA	3900
CCTGTCTCTG	TCTAAAGAGG	TGACCCCGGA	AGTGTTCCTG	GCACTGCGAG	CCAAGCCTGGA	3960
CCCAAGAGTCC	AGACACCCAC	CTGTGCCCCC	GCTTCTGCGG	TCTACCAAGGA	ACCCTCTAGG	4020
CCCAAGAGGGG	ACTTCTCTGT	TGCGCTTGGG	TGGAAGAGAGG	CCTCCTATTG	TCCTCGTAGA	4080
GGAAGCCACC	CGCGGGGCTG	AGGATGAGCC	AAGTGGGATT	CGGGGAACCG	CGTGGCTGGG	4140
GCGCCAGCCCC	GCGCTGCGTG	GCTGTCATGC	CTCCTGTATA	AGCGCCCAAG	CCTGCTGTCT	4200
CAGCCCTCCA	CTCCCTGCGAG	AGCTCAGAGG	CACGACCCCA	GCGATATCCC	TGCAAGCCATG	4260
AAGTGCGCTC	TGCTTGCGCT	GCGCGTGGCC	CTGCGCTGTG	GCGTCCAGGG	CATCATCGTG	4320
ACCCAGACCA	TGAAAGGGCT	GACATCCAG	AAAGTTGCGAG	GTTTGCGCGG	GTGGGTGAGT	4380
TGCAAGGGCG	GCAAGGGAGC	TGGCGCTCAG	AGAGCCAAAG	GAGGCTGTGA	CGTTGGGTTT	4440
CCATCAAGTCA	GCTAGGGGCA	CCTGACAAAT	CCCCGCTGGG	GCAAGCTTCA	CCAAGGCGTTT	4500
ACTGTCTTG	ATTCTGAGAG	CTGGAAGCGC	AAAGATCCAGG	TGTTGGCAGG	GCTGGCTTCT	4560
CCTGCGGGCG	CTCTCTGCGG	AGCAGACGGC	CGTCTTCTCC	AGTCTCTGCG	GCGCGCTGAT	4620
TTCTCTTTCC	TGTGAGGGCA	CCAAGGCTGC	TGGAAGCAGG	CCTGCGTGGG	CAGCTTCACA	4680
CGACCTTTGT	CATCTCTTTA	AAAGCCATGT	CTCCAAGAGTC	ATGTGTTGAA	GTTCTGCGGG	4740
TTAGTGGGAC	ACAAGTTCAAG	CCCTAAAAAG	GTCTCTCTGC	CCCTCAAAAT	TTCCCCACCT	4800
CCAAGCCATGT	CTCCCCAAAG	TCCAAATGTT	GCTACATGTG	GCGGGGGCTCA	TCTGGGTCCC	4860
TCTTTGGGTT	CAGTGTGAGT	CTGGGGAGAG	CATTCCCCAG	GGTGCAAGAGT	TGGGGGGAGT	4920
ATCTCAAGGG	TGCCCCAGGG	GCGGTGGGAC	AGAGAGGGCA	CTGTGGGGCT	GCGGGGGGCT	4980

-continued

TCCCACCCCC	A0A0T0CAAC	TCAA00TCCC	TCTCCA00T0	0C0000GACTT	00CACTCCTT	5040
00CTAT00C0	0CCA0C0ACA	TCTCCCT0CT	00AT0CCCCA0	A0T0CCCCC	T0A0A0T0TA	5100
C0T00A00A0	CT0AA0CCCA	CCCCC0A000	CAACCT00A0	ATCCT0CT0C	A0AAAT00T0	5160
00C0TCTCTC	CCCAACAT00	AACCCCCACT	CCCCA000CT	0T00ACCCCC	C000000T00	5220
00T0CA00A0	00ACCA000C	CCCCA000CT0	000AA0A000	CTCA0A0TTT	ACT00TACCC	5280
00C0CTCCAC	CCAA00CT0C	CCACCCA000	CITTTTTTTT	TTTTAAACTT	TTATTAATTT	5340
0AT0CTTCA0	AACATCATCA	AACAAAT0AA	CATAAAACAT	TCATTTTT0T	TTACTT00AA	5400
0000A0ATAA	AATCCTCT0A	A0T00AAAT0	CATA0CAAA0	ATACATACAA	T0A00CA00T	5460
ATTCT0AATT	CCCT0TTA0T	CT0A00ATTA	CAA0T0TATT	T0A0CAACAG	A0A0ACATTT	5520
TCATCATTTT	TA0TCT0AAC	ACCTCA0TAT	CTAAAAAT0AA	CAA0AA0TCC	T00AAAC0AA	5580
0CA0T0T000	0ATA00CCCC	T0T0AA00CT	0CT000A00C	A0CA0ACCT0	00TCTT0000	5640
CTCAA0CA0T	TCCC0CTACC	A0CCCT0TCC	ACCTCA0AC0	0000TCA000	T0CA00A0A0	5700
A0CT00AT00	0T0T00000C	A0A0AT0000	ACCT0AAACCC	CA000CT0CC	TTTT00000T	5760
0CCT0T00TC	AA00CTCTCC	CT0ACCTTTT	CTCTCT00CT	TCATCT0ACT	TCTCCT00CC	5820
CATCCACCC0	0TCCCCCT0T0	0CCT0A00T0	ACA0T0A0T0	C0CC0A00CT	A0TT00CCA0	5880
CT00CTCCTA	T0CCCAT0CC	ACCCCCCTCC	A0CCCTCCT0	00CCA0CTTC	T0CCCCCT00C	5940
CCTCA0TTCA	TCCT0AT0AA	AAT00TCCAT	0CCAAT00CT	CA0AAA0CA0	CT0TCTTTCA	6000
000A0AAC00	C0A0T0T0CT	CA0AA0AA0A	TTATT0CA0A	AAAAACCAAG	ATCCCT0C00	6060
T0TTCAA0AT	C0AT00T0A0	TCC000TCCC	T00000ACAC	CCACCACCCC	C0CCCCC000	6120
0ACT0T00AC	A00TTCA000	00CT00C0TC	000CCCT000	AT0CTAA000	ACT00T00T0	6180
AT0AA0ACAC	T0CCTT0ACA	CCT0CTTCAC	TT0CCTCCCC	T0CCACCT0C	CC0000CCTT	6240
0000C00T00	CCAT000CA0	0TCCCC00CT0	0C000CTAAC	CCACCA000T	0ACACCC0A0	6300
CTCTCTTT0C	T000000C00	0C00T0CTCT	000CCCTCA0	0CT0A0CTCA	00A00TACCT	6360
0T0CCCTCCC	A0000TAACC	0A0A0CC0TT	0CCCACTCCA	0000CCCCA00	T0CCCCAC0A	6420
CCCCA0CCCC	CTCCACA0CT	CCTTCATCTC	CT00A0ACAA	ACTCT0TCC0	CCCTC0CTCA	6480
TTCACTT0TT	C0TCCTAAAT	CC0A0AT0AT	AAA0CTTC0A	0000000TT0	000TTCCATC	6540
A000CT0CCC	TTCC0CC000	CA0CCT000C	CACATCT0CC	CTT00CCCCC	TCAG0ACTCA	6600
CTCT0ACT00	A00CCCT0CA	CT0ACT0AC0	CCA000T0CC	CA0CCCA000	TCTCT00C0C	6660
CATCCA0CT0	CACT000TTT	000T0CT00T	CCT0CCCCCA	A0CT0CCCC0	ACACCACA00	6720
CA0CC0000C	T0CCCACT00	CCTC00TCA0	00T0A0CCCC	A0CT0CCCCC	0CTCA000CT	6780
T0CCCC0ACA	AT0ACCCCAT	CCTCA00AC0	CACCCCCCTT	CCCTT0CT00	0CA0T0TCCA	6840
0CCCCACCC0	A0ATC00000	AA0CCCTATT	TCTT0ACAAC	TCCA0TCCCT	00000A0000	6900
0CCTCA0ACT	0A0T00T0A0	T0TTCCCAAG	TCCA00A00T	00T00A000T	CCT00C00AT	6960
CCA0A0TT0A	CA0T0A000C	TTCCCT000CC	CCAT0C0CCT	00CA0T00CA	0CA000A00A	7020
00AA0CACCA	TTTCA0000T	00000AT0CC	A0A00C0CTC	CCCACCC0T	CTTC0CC000	7080
T00T0ACCCC	00000A0CCC	C0CT00T0CT	00A000T0CT	00000CT0AC	TA0CAACCCC	7140
TCCCCCCCC0	TT00AACTCA	CTTTTCTCCC	0TCTT0ACCC	C0TCCA0CCT	T0AAT0A0AA	7200
CAAA0TCCIT	0T0CT00ACA	CC0ACTACAA	AAA0TACCT0	CTCTTCT0CA	T00AAAACAG	7260
T0CT0A0CCC	0A0CAAA0CC	T00CCT0CCA	0T0CCT000T	000T0CCAAC	CCT00CT0CC	7320
CA000A0ACC	A0CT0C0T00	TCCTT0CT0C	AACA00000T	000000T000	A0CTT0ATCC	7380

-continued

CCAAGAGAG	GAGGAGTGG	GGTCCCTGA	GTCCGCCAG	GAGAGAGTGG	TCCATACCG	7440
GAAGCCAATC	TCTGTGGG	CTGTGGGTGG	CTGGGGACGG	GGGCCAGACA	CACAGGCCGG	7500
GAGACGGGTG	GGCTGAGAA	CTGTGACTGG	TGTGACCCTC	GCATGGGGG	CGGTGGTCAC	7560
TGAATCTAAC	AGCCTTTGTT	ACCAGGAGT	TTCAATTATT	TCCCAAAATA	AGAACTCAAG	7620
TACAAAGCCA	TCTTTCAACT	ATCACATCCT	GAAAACAAAT	GGCAGGTGAC	ATTTTCTGTG	7680
CCGTAGCAGT	CCCACTGGG	ATTTTCAAGG	CCCCTGTGCC	AGGGGGGGCG	GGGCATCGGC	7740
GAGTGGAGGG	TCTGTGGCTGT	GTCAAGCCGG	CCAAGGGGAG	GAAAGGACCC	GGACAGCCAG	7800
AGGTGGGGGG	CAGGCTTTCC	CCCTGTGACC	TGCAAGCCCA	CTGCACTGCC	CTGGGAGGAA	7860
GGGAGGGGAA	CTAGGCCAAG	GGGAGAGGG	AGGTGCTCTG	GAGGGCAAGG	GCAAGACCTGC	7920
AGACCACCTT	GGGAGAGCAG	GACTGACCCC	CGTCCCTGCC	CCATAGTCAAG	GACCCCGGAG	7980
GTGGACAACG	AGGCCCTGGA	GAAATTGAC	AAAAGCCCTCA	AGGCCCTGCC	CATGCACATC	8040
CGGCTTGCTT	TCAACCCGAC	CCAAGTGGAG	GGTGAAGCACC	CAGGCCCGCG	CCTTCCCCAG	8100
GGCAGGAGGG	ACCCGGGGGG	GGGAGGAGCT	CCTCCCATGG	TGACCCCGAG	CTCCCCAGGG	8160
CTCCCCAGAG	GAAAGGGGTGG	GGTGAAGCAG	CCCCTGGGGG	CCCCCTCCCC	ACCCCGTGGC	8220
AGGCCCTCTT	TCCCGAGGTT	TCCAAGTCCA	TGCTGACCCC	CCCATGACTC	TCCCTCCCCC	8280
ACAGGGCAGT	GCCACGTCTA	GGTGAAGCCCC	TGCCGGTGGC	TCTGGGGTAA	GCTGCGTGGC	8340
CTGCCCCAGG	TGCTGGGAC	ACACATGGGG	TAGGGGGTCT	TGGTGGGGGG	TGGGACCCCA	8400
CATCAGGGGG	TGGGGTCCCC	CCTGTGAGAA	TGGCTGGGAG	CTGGGGTCCC	TGCTGGGAGC	8460
TGCAAGAGCTG	GCTGGGGCGG	TGCCACTCTT	GTGGGTGACC	TGTGTCTGGG	CCTCACACAC	8520
TGACCTCCTC	CAGCTCCTTC	CAGCAGAGCT	AAAGCTAAGT	GAGCCAGAAAT	GGTACCTAAG	8580
GGGAGGGCTAG	CGGTCTTTCT	CCCGAGGAGG	GGCTGTCTGT	GAAACCACCA	CCATGGAGAG	8640
GCTGGCAAGG	GTCTGGCAGG	TGCCCCAGGA	ATCACAAGGG	GGCCCCATGT	CCATTTCAAG	8700
GGCCGGGAGC	CTTGGACTCC	TCTGGGGACA	GACGACGTCA	CCACCGGGGG	CCCCCATCA	8760
GGGGGACTAG	AAAGGACCA	GACTGCAATC	ACCCTTCTGT	GGACCCAGGG	CCCTCCAGGG	8820
CCCTCCTGGG	GCTCCTGTCT	TGGGAGGCTT	CTCCTTCACC	AATAAGGCA	TAAACCTGTG	8880
CTCTCCCTTC	TGAGTCTTTG	CTGGAGGAGG	GGCAGGGGGT	GGAGAGAGTGG	TGGGGAGGGA	8940
GTCTGGCTCA	GAGGATGACA	GGGGGGCTGG	GATCCAAGGG	GTCTGATCA	CAGTCTTTGT	9000
ACAACTGGGG	GGCCACACAC	ATCACTGGGG	CTCTTTGAAA	CTTTCAAGGA	CCAGGGAGGG	9060
ACTGGGAGG	GACATCTGGC	AGTTCACTTG	GAGTGTTCAG	TCAACACCCA	AACTGACAA	9120
AGGACAGAAA	GTGGAGAAATG	GCTGTCTCTT	AGTCTAATAA	ATATTGATAT	GAAACTCAAG	9180
TTGCTCATGG	ATCAATATGG	CTTTATGATC	CAGCCAGCCA	CTACTGTGGT	ATCAACTCAT	9240
GTACCCAAAC	GCACTGATCT	GTCTGGCTAA	TGATGAGAGG	TTCCCAAGTGG	AGAGCTGGCA	9300
AGAGGTGACA	GTGAGAACTG	TCTGACACAC	CAGCAGAGTC	CACCAATCAT	CCTAAGGAGG	9360
TCAATCCTGG	TGTTCAATGG	AGGACTGATG	TGAGAGCTGA	AACTCCAATG	CTTTGGCCAC	9420
CTGATGTGAA	GAGCTGACTC	ATTTGAAAAA	ACCCGTGATG	TGGGAGAGAT	TGAGGGCAGG	9480
AGGAGAGAGG	GACGACAGAG	GATGAGATGG	TTGGATGGCA	TACCAACAC	AATGGACATG	9540
GGTTTGGGGT	GACTCCAAGG	GTGGGTGATG	GACAGGGAGG	CCTGGCGTGG	TACGGAGGCG	9600
GTTTATGGGG	TCACAAAGAC	TGAGTGAATG	AACTGAGCTG	AACTGAGATG	AAATGAGGTA	9660
TACAGCAAGG	TGGGGATTTT	TTAGATAATA	AGAAATATAC	CATAACATAG	TGTATACTCA	9720
TATTTTATAT	CATACCTGAA	TGCTCAATCA	CTCAATCTGA	TCTGACTCTG	TGACCTATGG	9780

-continued

ACCCTAGCCT	TCCAAGTTTC	TTCTGTCCAC	AGAAATTCCTC	AAAGCAAGAA	TACTGGAATG	9840
GCTAGCCATT	TCCTCCTCCA	GGGATCCTC	CCGACCCAGG	GATTGAACCG	GCATCTCCTG	9900
TATTGCAAG	TGGATTCTTT	ACCACTGTGC	CACCAAGGAA	GCCCCTGTTA	CTCTCTATGT	9960
CCCACCTAAT	TACCAAAAGCT	GCTCCAAAGAA	AAAGCCCCCTG	TCCCCCTCTGA	GCTTCCCCGCG	10020
CTGCAGAGGG	TGGTGGGGGT	AGACTGTGAC	CTGGGAACAC	CCTCCCCTTT	CAGGACTCCC	10080
GGGCCACGTG	ACCCACAGTC	CTGCAGACAG	CCGGTAAGCT	CTGCTCTTCA	AGGCTCATTG	10140
TCTTTAAAAA	AAACTGAGGT	CTATTTTGTG	ACTTCGCTGC	CGTAACTTCT	GAACATCCAAG	10200
TGCGATGAG	AGGACCTCCT	CCCCAGGCGT	CAGGGGCTTC	AGGGAAGCCA	CCTTCACCTA	10260
TGAGTCACCA	GACACTCGGG	GGTGGCCCCG	CCTTCAGGCT	GCTCACAGTC	TTCCCATCGT	10320
CCTGATCAAA	GAOCAAGACC	AATGACTTCT	TAGGAAGCAA	CAGACACCCA	CAGGACACTG	10380
AGGTTACCA	GAAGTGAAGT	GTCCTTTTGA	ACCTAAAGAC	ACACAAGCTCT	CGAAGGTTTT	10440
CTCTTTAATC	TGGATTTAAG	GCCTACTTGC	CCCTCAAGAG	GGAAAGACAGT	CCTGCATGTC	10500
CCCCAGACAG	CCACTCGGTG	GCATCCGAGG	CCACTTAAGTA	TTATCTGACC	GCACCCTGGA	10560
ATTAATCGGT	CCAAACTGGA	CAAAAACCTT	GGTGGGAAGT	TTCATCCCA	AGGCCTCAAC	10620
CATCCTGCTT	TGACCACCCT	GCATCTTTTT	TTCTTTTATG	TGTATGCATG	TATATATATA	10680
TATATATTTT	TTTTTTTTTC	ATTTTTTGGC	TGTGCTGGCT	GTTCGTTGCA	GTTCGGTGGC	10740
CAGGCTTCTC	TCTAGTTTCT	CTCTAGTCTT	CTCTTATCAC	AGAGCAATCT	CTAGACGATC	10800
GACGCGT						10807

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AATTCCGATC GACGCCTCGA CGATATACTC TAGACGATCG ACGCCTA

47

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: BLGAMP3

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TGGATCCCCCT GCCGCTGCTT CTGG

24

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: BLGAMP4

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AACGCGTCAT CCTCTGTGAG CCAG

24

G65710-3343660

-continued

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6839

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACTACOTAGT

10

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 42 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6632

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGACGCGGAT CCTACGTACC TGCAGCCATG TTTTCCATGA GG

42

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6627

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGGCTTCGG CAAGCTTCAG G

21

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6521

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GCCAAAGACT TACTTCCTC TAGA

24

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6520

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCATGAACGT CCGTGGTGG TTGTGCTACC

30

G C C T T G G C A A G C T T C A G G

-continued

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 30 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6519

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ACCACGCGAC GTTCATGCTC TAAAAACGTT

3 0

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6518

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCTGCGGGAT CCTACGTACT ACGGGGACAG GGAAGG

3 6

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6629

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAAAGGATGG TTTCT

4 5

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 45 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6630

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGACGCGAAT TCTACGTACC TGCAGCCATG AAACATCTAT TATTO

4 5

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (v i i) IMMEDIATE SOURCE:
 (B) CLONE: ZC6625

- (x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTGAAGATTT CAGATCTTGT C

2 1

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:
(B) CLONE: ZC6626

(x i) SEQUENCE DESCRIPTION: SBQ ID NO-21:

AAGAATTACT GTGGCCTACC A

21

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:
(B) CLONE: Z06624

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GCTGCGGAAT TCTACGTACT ATTGCTGTGG GAA

33

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:
(B) CLONE: ZC6314

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

COACOCOOAT CCTACOTACC TOCAOCCATO AOTTOOTCCT TOCAC

45

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:
(B) CLONE: #6517

(i i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTCTCTGGTA OCAACATACT A

21

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:
(B) CLONE: no6516

(x i) SEQUENCE DESCRIPTION: SBQ ID NO-23:

000TTTCTAG CCCTACTAOT AG

22

[illegible]

-continued

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(v i i) IMMEDIATE SOURCE:

- (B) CLONE: sc6515

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGGTTTCTAG CCCTACTAGT AG

22

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

AAGCTACGG TCGATCGTCT AGAGTATATC GTCGACGCGT CGATCGG

47

GAGTTCGTCGATCGTCTAGAGTATATCGTCGACGCGTCCGATCGG